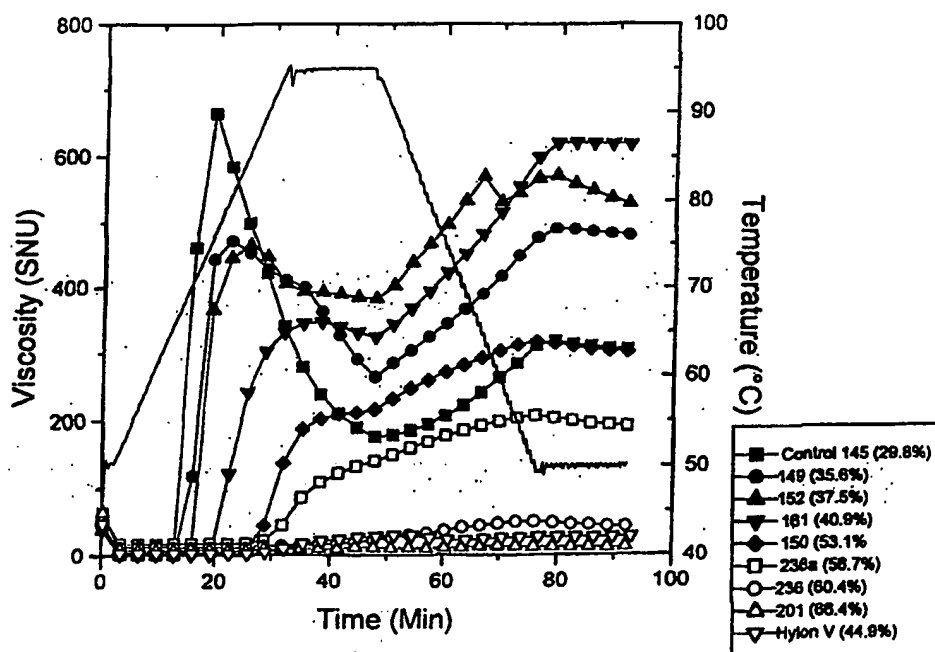




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification: C12N 15/82, 15/54, C08B 30/04, A01H 5/00		(11) International Publication Number: WO 96/34968
A2		(43) International Publication Date: 7 November 1996 (07.11.96)
(21) International Application Number: PCT/GB96/01075		[GB/GB]; 46 Castle Street, Wellingborough, Northamptonshire NN8 1LW (GB).
(22) International Filing Date: 3 May 1996 (03.05.96)		(74) Agent: KEITH W. NASH & CO. ; 90-92 Regent Street, Cambridge CB2 1DP (GB).
(30) Priority Data: 9509229.2 5 May 1995 (05.05.95) GB 9607409.1 10 April 1996 (10.04.96) GB		(81) Designated States: AU, BR, CA, JP, KR, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
(71) Applicant (for all designated States except US): NATIONAL STARCH AND CHEMICAL INVESTMENT HOLDING CORPORATION [US/US]; Suite 27, 501 Silverside Road, Wilmington, DE 19809 (US).		Published Without international search report and to be republished upon receipt of that report.
(72) Inventors; and		
(75) Inventors/Applicants (for US only): COOKE, David [GB/GB]; 2 Hunts Path, Oakley, Bedfordshire MK43 7SR (GB). DEBET, Martine [FR/GB]; 12 Bentley Close, Rectory Farm, Northampton NN3 5JS (GB). GIDLEY, Michael, John [GB/GB]; 45 Holmes Avenue, Raunds, Northamptonshire NN9 6SZ (GB). JOBLING, Stephen, Alan [GB/GB]; 19 Burwell Road, Eaton Socon, Huntingdon PE19 3QQ (GB). SAFFORD, Richard [GB/GB]; 10 Furness Close, Bedford, Bedfordshire MK41 8RN (GB). SIDEBOTTOM, Christopher, Michael [GB/GB]; 35 Waterloo Road, Bedford, Bedfordshire MK40 3PQ (GB). WESTCOTT, Roger, John		

(54) Title: IMPROVEMENTS IN OR RELATING TO PLANT STARCH COMPOSITION



(57) Abstract

Disclosed is a nucleotide sequence encoding an effective portion of a class A starch branching enzyme (SBE) obtainable from potato plants, a functional equivalent thereof, together with, inter alia, a corresponding polypeptide, a method of altering the characteristics of a plant, a plant having altered characteristics; and starch, particularly starch obtained from a potato plant, having novel properties.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

Title: Improvements in or Relating to Plant Starch Composition

Field of the Invention

This invention relates to novel nucleotide sequences, polypeptides encoded thereby, vectors and host cells and host organisms comprising one or more of the novel sequences, and to a method of altering one or more characteristics of an organism. The invention also relates to starch having novel properties and to uses thereof.

Background of the Invention

Starch is the major form of carbon reserve in plants, constituting 50% or more of the dry weight of many storage organs - e.g. tubers, seeds of cereals. Starch is used in numerous food and industrial applications. In many cases, however, it is necessary to modify the native starches, via chemical or physical means, in order to produce distinct properties to suit particular applications. It would be highly desirable to be able to produce starches with the required properties directly in the plant, thereby removing the need for additional modification. To achieve this via genetic engineering requires knowledge of the metabolic pathway of starch biosynthesis. This includes characterisation of genes and encoded gene products which catalyse the synthesis of starch. Knowledge about the regulation of starch biosynthesis raises the possibility of "re-programming" biosynthetic pathways to create starches with novel properties that could have new commercial applications.

The commercially useful properties of starch derive from the ability of the native granular form to swell and absorb water upon suitable treatment. Usually heat is required to cause granules to swell in a process known as gelatinisation, which has been defined (W A Atwell *et al*, Cereal Foods World 33, 306-311, 1988) as "... the collapse (disruption) of molecular orders within the starch granule manifested in irreversible changes in properties such as granular swelling, native crystallite melting, loss of birefringence, and starch solubilisation. The point of initial gelatinisation and the range over which it occurs is governed by starch concentration, method of observation, granule type, and heterogeneities within the granule population under observation". A number of techniques are available

for the determination of gelatinisation as induced by heating, a convenient and accurate method being differential scanning calorimetry, which detects the temperature range and enthalpy associated with the collapse of molecular orders within the granule. To obtain accurate and meaningful results, the peak and/or onset temperature of the endotherm observed by differential scanning calorimetry is usually determined.

The consequence of the collapse of molecular orders within starch granules is that the granules are capable of taking up water in a process known as pasting, which has been defined (W A Atwell *et al*, Cereal Foods World 33, 306-311, 1988) as "*... the phenomenon following gelatinisation in the dissolution of starch. It involves granular swelling, exudation of molecular components from the granule, and eventually, total disruption of the granules*". The best method of evaluating pasting properties is considered to be the viscoamylograph (Atwell *et al*, 1988 cited above) in which the viscosity of a stirred starch suspension is monitored under a defined time/temperature regime. A typical viscoamylograph profile for potato starch shows an initial rise in viscosity, which is considered to be due to granule swelling. In addition to the overall shape of the viscosity response in a viscoamylograph, a convenient quantitative measure is the temperature of initial viscosity development (onset). Figure 1 shows such a typical viscosity profile for potato starch, during and after cooking, and includes stages A-D which correspond to viscosity onset (A), maximum viscosity (B), complete dispersion (C) and reassociation of molecules (or retrogradation, D). In the figure, the dotted line represents viscosity (in stirring number units) of a 10% w/w starch suspension and the unbroken line shows the temperature in degrees centigrade. At a certain point, defined by the viscosity peak, granule swelling is so extensive that the resulting highly expanded structures are susceptible to mechanically-induced fragmentation under the stirring conditions used. With increased heating and holding at 95°C, further reduction in viscosity is observed due to increased fragmentation of swollen granules. This general profile has previously always been found for native potato starch.

After heating starches in water to 95°C and holding at that temperature (for typically 15 minutes), subsequent cooling to 50°C results in an increase in viscosity due to the process of retrogradation or set-back. Retrogradation (or set-back) is defined (Atwell *et al.*, 1988

cited above) as "...a process which occurs when the molecules comprising gelatinised starch begin to reassociate in an ordered structure...". At 50°C, it is primarily the amylose component which reassociates, as indicated by the increase in viscoamylograph viscosity for starch from normal maize (21.6% amylose) compared with starch from waxy maize (1.1% amylose) as shown in Figure 2. Figure 2 is a viscoamylograph of 10%w/w starch suspensions from waxy maize (solid line), conventional maize (dots and dashes), high amylose variety (hylon 5, dotted line) and a very high amylose variety (hylon 7, crosses). The temperature profile is also shown by a solid line, as in Figure 1. The extent of viscosity increase in the viscoamylograph on cooling and holding at 50°C depends on the amount of amylose which is able to reassociate due to its exudation from starch granules during the gelatinisation and pasting processes. A characteristic of amylose-rich starches from maize plants is that very little amylose is exuded from granules by gelatinisation and pasting up to 95°C, probably due to the restricted swelling of the granules. This is illustrated in Figure 2 which shows low viscosities for a high amylose (44.9%) starch (Hylon 5) from maize during gelatinisation and pasting at 95°C and little increase in viscosity on cooling and holding at 50°C. This effect is more extreme for a higher amylose content (58%, as in Hylon 7), which shows even lower viscosities in the viscoamylograph test (Figure 2). For commercially-available high amylose starches (currently available from maize plants, such as those described above), processing at greater than 100°C is usually necessary in order to generate the benefits of high amylose contents with respect to increased rates and strengths of reassociation, but use of such high temperatures is energetically unfavourable and costly. Accordingly, there is an unmet need for starches of high amylose content which can be processed below 100°C and still show enhanced levels of reassociation, as indicated for example by viscoamylograph measurements.

The properties of potato starch are useful in a variety of both food and non-food (paper, textiles, adhesives etc.) applications. However, for many applications, properties are not optimum and various chemical and physical modifications well known in the art are undertaken in order to improve useful properties. Two types of property manipulation which would be of use are: the controlled alteration of gelatinisation and pasting temperatures; and starches which suffer less granular fragmentation during pasting than

conventional starches.

Currently the only ways of manipulating the gelatinisation and pasting temperatures of potato starch are by the inclusion of additives such as sugars, polyhydroxy compounds of salts (Evans & Haisman, *Starke* 34, 224-231, 1982) or by extensive physical or chemical pre-treatments (e.g. Stute, *Starke* 44, 205-214, 1992). The reduction of granule fragmentation during pasting can be achieved either by extensive physical pretreatments (Stute, *Starke* 44, 205-214, 1992) or by chemical cross-linking. Such processes are inconvenient and inefficient. It is therefore desirable to obtain plants which produce starch which intrinsically possesses such advantageous properties.

Starch consists of two main polysaccharides, amylose and amylopectin. Amylose is a generally linear polymer containing α -1,4 linked glucose units, while amylopectin is a highly branched polymer consisting of a α -1,4 linked glucan backbone with α -1,6 linked glucan branches. In most plant storage reserves amylopectin constitutes about 75% of the starch content. Amylopectin is synthesized by the concerted action of soluble starch synthase and starch branching enzyme [α -1,4 glucan: α -1,4 glucan 6-glycosyltransferase, EC 2.4.1.18]. Starch branching enzyme (SBE) hydrolyses α -1,4 linkages and rejoins the cleaved glucan, via an α -1,6 linkage, to an acceptor chain to produce a branched structure. The physical properties of starch are strongly affected by the relative abundance of amylose and amylopectin, and SBE is therefore a crucial enzyme in determining both the quantity and quality of starches produced in plant systems.

In most plants studied to date e.g. maize (Boyer & Preiss, 1978 *Biochem. Biophys. Res. Comm.* 80, 169-175), rice (Smyth, 1988 *Plant Sci.* 57, 1-8) and pea (Smith, *Planta* 175, 270-279), two forms of SBE have been identified, each encoded by a separate gene. A recent review by Burton *et al.*, (1995 *The Plant Journal* 7, 3-15) has demonstrated that the two forms of SBE constitute distinct classes of the enzyme such that, in general, enzymes of the same class from different plants may exhibit greater similarity than enzymes of different classes from the same plant. In their review, Burton *et al.* termed the two respective enzyme families class "A" and class "B", and the reader is referred thereto (and to the references cited therein) for a detailed discussion of the distinctions

between the two classes. One general distinction of note would appear to be the presence, in class A SBE molecules, of a flexible N-terminal domain, which is not found in class B molecules. The distinctions noted by Burton *et al.* are relied on herein to define class A and class B SBE molecules, which terms are to be interpreted accordingly.

However in potato, only one isoform of the SBE molecule (belonging to class B) has thus far been reported and only one gene cloned (Blennow & Johansson, 1991 *Phytochem.* 30, 437-444, and Koßmann *et al.*, 1991 *Mol. Gen. Genet.* 230, 39-44). Further, published attempts to modify the properties of starch in potato plants (by preventing expression of the single known SBE) have generally not succeeded (e.g. Müller-Rober & Koßmann 1994 *Plant Cell and Environment* 17, 601-613).

Summary of the Invention

In a first aspect the invention provides a nucleotide sequence encoding an effective portion of a class A starch branching enzyme (SBE) obtainable from potato plants.

Preferably the nucleotide sequence encodes a polypeptide comprising an effective portion of the amino acid sequence shown in Figure 5 (excluding the sequence MNKRIDL, which does not represent part of the SBE molecule), or a functional equivalent thereof (which term is discussed below). The amino acid sequence shown in Figure 5 (Seq ID No. 15) includes a leader sequence which directs the polypeptide, when synthesised in potato cells, to the amyloplast. Those skilled in the art will recognise that the leader sequence is removed to produce a mature enzyme and that the leader sequence is therefore not essential for enzyme activity. Accordingly, an "effective portion" of the polypeptide is one which possesses sufficient SBE activity to complement the branching enzyme mutation in *E. coli* KV 832 cells (described below) and which is active when expressed in *E. coli* in the phosphorylation stimulation assay. An example of an incomplete polypeptide which nevertheless constitutes an "effective portion" is the mature enzyme lacking the leader sequence. By analogy with the pea class A SBE sequence, the potato class A sequence shown in Figure 5 probably possesses a leader sequence of about 48 amino acid residues, such that the N terminal amino acid sequence is thought to commence around the glutamic acid residue (E) at position 49 (EKSSYN... etc.). Those skilled in the art will appreciate

that an effective portion of the enzyme may well omit other parts of the sequence shown in the figure without substantial detrimental effect. For example, the C-terminal glutamic acid-rich region could be reduced in length, or possibly deleted entirely, without abolishing class A SBE activity. A comparison with other known SBE sequences, especially other class A SBE sequences (see for example, Burton *et al*, 1995 cited above), should indicate those portions which are highly conserved (and thus likely to be essential for activity) and those portions which are less well conserved (and thus are more likely to tolerate sequence changes without substantial loss of enzyme activity).

Conveniently the nucleotide sequence will comprise substantially nucleotides 289 to 2790 of the DNA sequence (Seq ID No. 14) shown in Figure 5 (which nucleotides encode the mature enzyme) or a functional equivalent thereof, and may also include further nucleotides at the 5' or 3' end. For example, for ease of expression, the sequence will desirably also comprise an in-frame ATG start codon, and may also encode a leader sequence. Thus, in one embodiment, the sequence further comprises nucleotides 145 to 288 of the sequence shown in Figure 5. Other embodiments are nucleotides 228 to 2855 of the sequence labelled "psbe2con.seq" in Figure 8, and nucleotides 57 to 2564 of the sequence shown in Figure 12 (preferably comprising an in-frame ATG start codon, such as the sequence of nucleotides 24 to 56 in the same Figure), or functional equivalents of the aforesaid sequences.

The term "functional equivalent" as applied herein to nucleotide sequences is intended to encompass those sequences which differ in their nucleotide composition to that shown in Figure 5 but which, by virtue of the degeneracy of the genetic code, encode polypeptides having identical or substantially identical amino acid sequences. It is intended that the term should also apply to sequences which are sufficiently homologous to the sequence of the invention that they can hybridise to the complement thereof under stringent hybridisation conditions - such equivalents will preferably possess at least 85%, more preferably at least 90%, and most preferably at least 95% sequence homology with the sequence of the invention as exemplified by nucleotides 289 to 2790 of the DNA sequence shown in Figure 5. It will be apparent to those skilled in the art that the nucleotide sequence of the invention may also find useful application when present as an "antisense"

sequence. Accordingly, functionally equivalent sequences will also include those sequences which can hybridise, under stringent hybridisation conditions, to the sequence of the invention (rather than the complement thereof). Such "antisense" equivalents will preferably possess at least 85%, more preferably at least 90%, and most preferably 95% sequence homology with the complement of the sequence of the invention as exemplified by nucleotides 289 to 2790 of the DNA sequence shown in Figure 5. Particular functional equivalents are shown, for example, in Figures 8 and 10 (if one disregards the various frameshift mutations noted therein).

The invention also provides vectors, particularly expression vectors, comprising the nucleotide sequence of the invention. The vector will typically comprise a promoter and one or more regulatory signals of the type well known to those skilled in the art. The invention also includes provision of cells transformed (which term encompasses transduction and transfection) with a vector comprising the nucleotide sequence of the invention.

The invention further provides a class A SBE polypeptide, obtainable from potato plants. In particular the invention provides the polypeptide in substantially pure form, especially in a form free from other plant-derived (especially potato plant-derived) components, which can be readily accomplished by expression of the relevant nucleotide sequence in a suitable non-plant host (such as any one of the yeast strains routinely used for expression purposes, e.g. *Pichia spp.* or *Saccharomyces spp.*). Typically the enzyme will substantially comprise the sequence of amino acid residues 49 to 882 shown in Figure 5 (disregarding the sequence MNKRIDL, which is not part of the enzyme), or a functional equivalent thereof. The polypeptide of the invention may be used in a method of modifying starch *in vitro*, comprising treating starch under suitable conditions (e.g. appropriate temperature, pH, etc) with an effective amount of the polypeptide according to the invention.

The term "functional equivalent", as applied herein to amino acid sequences, is intended to encompass amino acid sequences substantially similar to that shown in Figure 5, such that the polypeptide possesses sufficient activity to complement the branching enzyme mutation in *E. coli* KV 832 cells (described below) and which is active in *E. coli* in the

phosphorylation stimulation assay. Typically such functionally equivalent amino acid sequences will preferably possess at least 85%, more preferably at least 90%, and most preferably at least 95% sequence identity with the amino acid sequence of the mature enzyme (i.e. minus leader sequence) shown in Figure 5. Those skilled in the art will appreciate that conservative substitutions may be made generally throughout the molecule without substantially affecting the activity of the enzyme. Moreover, some non-conservative substitutions may be tolerated, especially in the less highly conserved regions of the molecule. Such substitutions may be made, for example, to modify slightly the activity of the enzyme. The polypeptide may, if desired, include a leader sequence, such as that exemplified by residues 1 to 48 of the amino acid sequence shown in Figure 5, although other leader sequences and signal peptides and the like are known and may be included.

A portion of the nucleotide sequence of the invention has been introduced into a plant and found to affect the characteristics of the plant. In particular, introduction of the sequence of the invention, operably linked in the antisense orientation to a suitable promoter, was found to reduce the amount of branched starch molecules in the plant. Additionally, it has recently been demonstrated in other experimental systems that "sense suppression" can also occur (i.e. expression of an introduced sequence operably linked in the sense orientation can interfere, by some unknown mechanism, with the expression of the native gene), as described by Matzke & Matzke (1995 Plant Physiol. 107, 679-685). Any one of the methods mentioned by Matzke & Matzke could, in theory, be used to affect the expression in a host of a homologous SBE gene.

It is believed that antisense methods are mainly operable by the production of antisense mRNA which hybridises to the sense mRNA, preventing its translation into functional polypeptide, possibly by causing the hybrid RNA to be degraded (e.g. Sheehy *et al.*, 1988 PNAS 85, 8805-8809; Van der Krol *et al.*, Mol. Gen. Genet. 220, 204-212). Sense suppression also requires homology between the introduced sequence and the target gene, but the exact mechanism is unclear. It is apparent however that, in relation to both antisense and sense suppression, neither a full length nucleotide sequence, nor a "native" sequence is essential. Preferably the "effective portion" used in the method will comprise

at least one third of the full length sequence, but by simple trial and error other fragments (smaller or larger) may be found which are functional in altering the characteristics of the plant.

Thus, in a further aspect the invention provides a method of altering the characteristics of a plant, comprising introducing into the plant an effective portion of the sequence of the invention operably linked to a suitable promoter active in the plant. Conveniently the sequence will be linked in the anti-sense orientation to the promoter. Preferably the plant is a potato plant. Conveniently, the characteristic altered relates to the starch content and/or starch composition of the plant (i.e. amount and/or type of starch present in the plant). Preferably the method of altering the characteristics of the plant will also comprise the introduction of one or more further sequences, in addition to an effective portion of the sequence of the invention. The introduced sequence of the invention and the one or more further sequences (which may be sense or antisense sequences) may be operably linked to a single promoter (which would ensure both sequences were transcribed at essentially the same time), or may be operably linked to separate promoters (which may be necessary for optimal expression). Where separate promoters are employed they may be identical to each other or different. Suitable promoters are well known to those skilled in the art and include both constitutive and inducible types. Examples include the CaMV 35S promoter (e.g. single or tandem repeat) and the patatin promoter. Advantageously the promoter will be tissue-specific. Desirably the promoter will cause expression of the operably linked sequence at substantial levels only in the tissue of the plant where starch synthesis and/or starch storage mainly occurs. Thus, for example, where the sequence is introduced into a potato plant, the operably linked promoter may be tuber-specific, such as the patatin promoter.

Desirably, for example, the method will also comprise the introduction of an effective portion of a sequence encoding a class B SBE, operably linked in the antisense orientation to a suitable promoter active in the plant. Desirably the further sequence will comprise an effective portion of the sequence encoding the potato class B SBE molecule. Conveniently the further sequence will comprise an effective portion of the sequence described by Blennow & Johansson (1991 *Phytochem.* 30, 437-444) or that disclosed in

WO92/11375. More preferably, the further sequence will comprise at least an effective portion of the sequence disclosed in International Patent Application No. WO 95/26407. Use of antisense sequences against both class A and class B SBE in combination has now been found by the present inventors to result in the production of starch having very greatly altered properties (see below). Those skilled in the art will appreciate the possibility that, if the plant already comprises a sense or antisense sequence which efficiently inhibits the class B SBE activity, introduction of a sense or antisense sequence to inhibit class A SBE activity (thereby producing a plant with inhibition of both class A and class B activity) might alter greatly the properties of the starch in the plant, without the need for introduction of one or more further sequences. Thus the sequence of the invention is conveniently introduced into plants already having low levels of class A and/or class B SBE activity, such that the inhibition resulting from the introduction of the sequence of the invention is likely to have a more pronounced effect.

The sequence of the invention, and the one or more further sequences if desired, can be introduced into the plant by any one of a number of well-known techniques (e.g. Agrobacterium-mediated transformation, or by "biolistic" methods). The sequences are likely to be most effective in inhibiting SBE activity in potato plants, but theoretically could be introduced into any plant. Desirable examples include pea, tomato, maize, wheat, rice, barley, sweet potato and cassava plants. Preferably the plant will comprise a natural gene encoding an SBE molecule which exhibits reasonable homology with the introduced nucleic acid sequence of the invention.

In another aspect, the invention provides a plant cell, or a plant or the progeny thereof, which has been altered by the method defined above. The progeny of the altered plant may be obtained, for example, by vegetative propagation, or by crossing the altered plant and reserving the seed so obtained. The invention also provides parts of the altered plant, such as storage organs. Conveniently, for example, the invention provides tubers comprising altered starch, said tubers being obtained from an altered plant or the progeny thereof. Potato tubers obtained from altered plants (or the progeny thereof) will be particularly useful materials in certain industrial applications and for the preparation and/or processing of foodstuffs and may be used, for example, to prepare low-fat waffles and

chips (amylose generally being used as a coating to prevent fat uptake), and to prepare mashed potato (especially "instant" mashed potato) having particular characteristics.

In particular relation to potato plants, the invention provides a potato plant or part thereof which, in its wild type possesses an effective SBE A gene, but which plant has been altered such that there is no effective expression of an SBE A polypeptide within the cells of at least part of the plant. The plant may have been altered by the method defined above, or may have been selected by conventional breeding to be deleted for the class A SBE gene, presence or absence of which can be readily determined by screening samples of the plants with a nucleic acid probe or antibody specific for the potato class A gene or gene product respectively.

The invention also provides starch extracted from a plant altered by the method defined above, or the progeny of such a plant, the starch having altered properties compared to starch extracted from equivalent, but unaltered, plants. The invention further provides a method of making altered starch, comprising altering a plant by the method defined above and extracting therefrom starch having altered properties compared to starch extracted from equivalent, but unaltered, plants. Use of nucleotide sequences in accordance with the invention has allowed the present inventors to produce potato starches having a wide variety of novel properties.

In particular the invention provides the following: a plant (especially a potato plant) altered by the method defined above, containing starch which, when extracted from the plant, has an elevated endotherm peak temperature as judged by DSC, compared to starch extracted from a similar, but unaltered, plant; a plant (especially a potato plant) altered by the method defined above, containing starch which, when extracted from the plant, has an elevated viscosity onset temperature (conveniently elevated by 10 - 25°C) as judged by viscoamylograph compared to starch extracted from a similar, but unaltered, plant; a plant (especially a potato plant) altered by the method defined above, containing starch which, when extracted from the plant, has a decreased peak viscosity (conveniently decreased by 240 - 700SNUs) as judged by viscoamylograph compared to starch extracted from a similar, but unaltered, plant; a plant (especially a potato plant) altered by the method

defined above, containing starch which, when extracted from the plant, has an increased pasting viscosity (conveniently increased by 37 - 260 SNUs) as judged by viscoamylograph compared to starch extracted from a similar, but unaltered, plant; a plant (especially a potato plant) altered by the method defined above, containing starch which, when extracted from the plant, has an increased set-back viscosity (conveniently increased by 224 - 313 SNUs) as judged by viscoamylograph compared to starch extracted from a similar, but unaltered, plant; a plant (especially a potato plant) altered by the method defined above, containing starch which, when extracted from the plant, has a decreased set-back viscosity as judged by viscoamylograph compared to starch extracted from a similar, but unaltered, plant; and a plant (especially a potato plant) altered by the method defined above, containing starch which, when extracted from the plant, has an elevated amylose content as judged by iodometric assay (i.e. by the method of Morrison & Laignelet 1983, cited above) compared to starch extracted from a similar, but unaltered, plant. The invention also provides for starch obtainable or obtained from such plants as aforesaid.

In particular the invention provides for starch which, as extracted from a potato plant by wet milling at ambient temperature, has one or more of the following properties, as judged by viscoamylograph analysis performed according to the conditions defined below:

viscosity onset temperature in the range 70-95°C (preferably 75-95°C); peak viscosity in the range 500 - 12 stirring number units; pasting viscosity in the range 214 - 434 stirring number units; set-back viscosity in the range 450 - 618 or 14 - 192 stirring number units; or displays no significant increase in viscosity during viscoamylograph. Peak, pasting and set-back viscosities are defined below. Viscosity onset temperature is the temperature at which there is a sudden, marked increase in viscosity from baseline levels during viscoamylograph, and is a term well-known to those skilled in the art.

In other particular embodiments, the invention provides starch which as extracted from a potato plant by wet milling at ambient temperature has a peak viscosity in the range 200 - 500 SNUs and a set-back viscosity in the range 275-618 SNUs as judged by viscoamylograph according to the protocol defined below; and starch which as extracted from a potato plant by wet milling at ambient temperature has a viscosity which does not decrease between the start of the heating phase (step 2) and the start of the final holding

phase (step 5) and has a set-back viscosity of 303 SNUs or less as judged by viscoamylograph according to the protocol defined below.

For the purposes of the present invention, viscoamylograph conditions are understood to pertain to analysis of a 10% (w/w) aqueous suspension of starch at atmospheric pressure, using a Newport Scientific Rapid Visco Analyser with a heating profile of: holding at 50°C for 2 minutes (step 1), heating from 50 to 95°C at a rate of 1.5°C per minute (step 2), holding at 95°C for 15 minutes (step 3), cooling from 95 to 50°C at a rate of 1.5°C per minute (step 4), and then holding at 50°C for 15 minutes (step 5). Peak viscosity may be defined for present purposes as the maximum viscosity attained during the heating phase (step 2) or the holding phase (step 3) of the viscoamylograph. Pasting viscosity may be defined as the viscosity attained by the starch suspensions at the end of the holding phase (step 3) of the viscoamylograph. Set-back viscosity may be defined as the viscosity of the starch suspension at the end of step 5 of the viscoamylograph.

In yet another aspect the invention provides starch from a potato plant having an apparent amylose content (% w/w) of at least 35%, as judged by iodometric assay according to the method described by Morrison & Laignelet (1983 J. Cereal Science 1, 9-20). Preferably the starch will have an amylose content of at least 40%, more preferably at least 50%, and most preferably at least 66%. Starch obtained directly from a potato plant and having such properties has not hitherto been produced. Indeed, as a result of the present invention, it is now possible to generate *in vivo* potato starch which has some properties analogous to the very high amylose starches (e.g. Hylon 7) obtainable from maize.

Starches with high (at least 35%) amylose contents find commercial application as, amongst other reasons, the amylose component of starch reassociates more strongly and rapidly than the amylopectin component during retrogradation processes. This may result, for example, in pastes with higher viscosities, gels of greater cohesion, or films of greater strength for starches with high (at least 35%) compared with normal (less than 35%) amylose contents. Alternatively, starches may be obtained with very high amylose contents, such that the granule structure is substantially preserved during heating, resulting in starch suspensions which demonstrate substantially no increase in viscosity during

cooking (i.e. there is no significant viscosity increase during viscoamylograph conditions defined above). Such starches typically exhibit a viscosity increase of less than 10% (preferably less than 5%) during viscoamylograph under the conditions defined above.

In commerce, these valuable properties are currently obtained from starches of high amylose content derived from maize plants. It would be of commercial value to have an alternative source of high amylose starches from potato as other characteristics such as granule size, organoleptic properties and textural qualities may distinguish application performances of high amylose starches from maize and potato plants.

Thus high amylose starch obtained by the method of the present invention may find application in many different technological fields, which may be broadly categorised into two groups: food products and processing; and "Industrial" applications. Under the heading of food products, the novel starches of the present invention may find application as, for example, films, barriers, coatings or gelling agents. In general, high amylose content starches absorb less fat during frying than starches with low amylose content, thus the high amylose content starches of the invention may be advantageously used in preparing low fat fried products (e.g. potato chips, crisps and the like). The novel starches may also be employed with advantage in preparing confectionery and in granular and retrograded "resistant" starches. "Resistant" starch is starch which is resistant to digestion by α -amylase. As such, resistant starch is not digested by α -amylases present in the human small intestine, but passes into the colon where it exhibits properties similar to soluble and insoluble dietary fibre. Resistant starch is thus of great benefit in foodstuffs due to its low calorific value and its high dietary fibre content. Resistant starch is formed by the retrogradation (akin to recrystallization) of amylose from starch gels. Such retrogradation is inhibited by amylopectin. Accordingly, the high amylose starches of the present invention are excellent starting materials for the preparation of resistant starch. Suitable methods for the preparation of resistant starch are well-known to those skilled in the art and include, for example, those described in US 5,051,271 and US 5,281,276. Conveniently the resistant starches provided by the present invention comprise at least 5% total dietary fibre, as judged by the method of Prosky *et al.*, (1985 J. Assoc. Off. Anal. Chem. 68, 677), mentioned in US 5,281, 276.

Under the heading of "Industrial" applications, the novel starches of the invention may be advantageously employed, for example, in corrugating adhesives, in biodegradable products such as loose fill packaging and foamed shapes, and in the production of glass fibers and textiles.

Those skilled in the art will appreciate that the novel starches of the invention may, if desired, be subjected *in vitro* to conventional enzymatic, physical and/or chemical modification, such as cross-linking, introduction of hydrophobic groups (e.g. octenyl succinic acid, dodecyl succinic acid), or derivatization (e.g. by means of esterification or etherification).

In yet another aspect the invention provides high (35% or more) amylose starches which generate paste viscosities greater than those obtained from high amylose starches from maize plants after processing at temperatures below 100°C. This provides the advantage of more economical starch gelatinisation and pasting treatments through the use of lower processing temperatures than are currently required for high amylose starches from maize plants.

The invention will now be further described by way of illustrative example and with reference to the drawings, of which:

Figure 1 shows a typical viscoamylograph for a 10% w/w suspension of potato starch;

Figure 2 shows viscoamylographs for 10% suspensions of starch from various maize varieties;

Figure 3 is a schematic representation of the cloning strategy used by the present inventors;

Figure 4a shows the amino acid alignment of the C-terminal portion of starch branching enzyme isoforms from various sources: amino acid residues matching the consensus

sequence are shaded;

Figure 4b shows the alignment of DNA sequences of various starch branching enzyme isoforms which encode a conserved amino acid sequence;

Figure 5 shows the DNA sequence (Seq ID No. 14) and predicted amino acid sequence (Seq ID No. 15) of a full length potato class A SBE cDNA clone obtained by PCR;

Figure 6 shows a comparison of the most highly conserved part of the amino acid sequences of potato class A (uppermost sequence) and class B (lowermost sequence) SBE molecules;

Figure 7 shows a comparison of the amino acid sequence of the full length potato class A (uppermost sequence) and pea (lowermost sequence) class A SBE molecules;

Figure 8 shows a DNA alignment of various full length potato class A SBE clones obtained by the inventors;

Figure 9 shows the DNA sequence of a potato class A SBE clone determined by direct sequencing of PCR products, together with the predicted amino acid sequence;

Figure 10 is a multiple DNA alignment of various full length potato SBE A clones obtained by the inventors;

Figure 11 is a schematic illustration of the plasmid pSJ64;

Figure 12 shows the DNA sequence and predicted amino acid sequence of the full length potato class A SBE clone as present in the plasmid pSJ90; and

Figure 13 shows viscoamylographs for 10% w/w suspensions of starch from various transgenic potato plants made by the relevant method aspect of the invention.

Examples

Example 1

Cloning of Potato class A SBE

The strategy for cloning the second form of starch branching enzyme from potato is shown in Figure 3. The small arrowheads represent primers used by the inventors in PCR and RACE protocols. The approximate size of the fragments isolated is indicated by the numerals on the right of the Figure. By way of explanation, a comparison of the amino acid sequences of several cloned plant starch branching enzymes (SBE) from maize (class A), pea (class A), maize (class B), rice (class B) and potato (class B), as well as human glycogen branching enzyme, allowed the inventors to identify a region in the carboxy-terminal one third of the protein which is almost completely conserved (GYLNFMGNEFGHPEWIDFPR) (Figure 4a). A multiple alignment of the DNA sequences (human, pea class A, potato class B, maize class B, maize class A and rice class B, respectively) corresponding to this region is shown in Figure 4b and was used to design an oligo which would potentially hybridize to all known plant starch branching enzymes: AATTT(C/T)ATGGGIAA(C/T)GA(A/G)TT(C/T)GG (Seq ID No. 20).

Library PCR

The initial isolation of a partial potato class A SBE cDNA clone was from an amplified potato tuber cDNA library in the λ Zap vector (Stratagene). One half μ L of a potato cDNA library (titre 2.3×10^8 pfu/mL) was used as template in a 50 μ L reaction containing 100 pmol of a 16 fold degenerate POTSBE primer and 25 pmol of a T7 primer (present in the λ Zap vector 3' to the cDNA sequences - see Figure 3), 100 μ M dNTPs, 2.5 U Taq polymerase and the buffer supplied with the Taq polymerase (Stratagene). All components except the enzyme were added to a 0.5 mL microcentrifuge tube, covered with mineral oil and incubated at 94°C for 7 minutes and then held at 55°C, while the Taq polymerase was added and mixed by pipetting. PCR was then performed by incubating for 1 min at 94°C, 1 min at 58°C and 3 minutes at 72°C, for 35 cycles. The PCR products were extracted with phenol/chloroform, ethanol precipitated and resuspended in TE pH 8.0 before cloning into the T/A cloning vector pT7BlueR (Invitrogen).

Several fragments between 600 and 1300 bp were amplified. These were isolated from an agarose gel and cloned into the pT7BlueR T/A cloning vector. Restriction mapping of 24 randomly selected clones showed that they belonged to several different groups (based on size and presence/absence of restriction sites). Initially four clones were chosen for sequencing. Of these four, two were found to correspond to the known potato class B SBE sequence, however the other two, although homologous, differed significantly and were more similar to the pea class A SBE sequence, suggesting that they belonged to the class A family of branching enzymes (Burton *et al.*, 1995 *The Plant Journal*, cited above). The latter two clones (~ 800bp) were sequenced fully. They both contained at the 5' end the sequence corresponding to the degenerate oligonucleotide used in the PCR and had a predicted open reading frame of 192 amino acids. The deduced amino acid sequence was highly homologous to that of the pea class A SBE.

The ~ 800 bp PCR derived cDNA fragment (corresponding to nucleotides 2281 to 3076 of the psbe2 con.seq sequence shown in Figure 8) was used as a probe to screen the potato tuber cDNA library. From one hundred and eighty thousand plaques, seven positives were obtained in the primary screen. PCR analysis showed that five of these clones were smaller than the original 800 bp cDNA clone, so these were not analysed further. The two other clones (designated 3.2.1 and 3.1.1) were approximately 1200 and 1500 bp in length respectively. These were sequenced from their 5' ends and the combined consensus sequence aligned with the sequence from the PCR generated clones. The cDNA clone 3.2.1 was excised from the phage vector and plasmid DNA was prepared and the insert fully sequenced. Several attempts to obtain longer clones from the library were unsuccessful, therefore clones containing the 5' end of the full length gene were obtained using RACE (rapid amplification of cDNA ends).

Rapid Amplification of cDNA ends (RACE) and PCR conditions

RACE was performed essentially according to Frohman (1992 *Amplifications* 11-15). Two μ g of total RNA from mature potato tubers was heated to 65°C for 5 min and quick cooled on ice. The RNA was then reverse transcribed in a 20 μ L reaction for 1 hour at 37°C using BRL's M-MLV reverse transcriptase and buffer with 1 mM DTT, 1 mM dNTPs, 1 U/ μ L RNasin (Promega) and 500 pmol random hexamers (Pharmacia) as

primer. Excess primers were removed on a Centricon 100 column and cDNA was recovered and precipitated with isopropanol. cDNA was A-tailed in a volume of 20 μ l using 10 units terminal transferase (BRL), 200 μ M dATP for 10 min at 37°C, followed by 5 min at 65°C. The reaction was then diluted to 0.5 ml with TE pH 8 and stored at 4°C as the cDNA pool. cDNA clones were isolated by PCR amplification using the primers $R_0R_1dT_{17}$, R_0 and POTSBE24. The PCR was performed in 50 μ L using a hot start technique: 10 μ L of the cDNA pool was heated to 94°C in water for 5 min with 25 pmol POTSBE24, 25 pmol R_0 and 2.5 pmol of $R_0R_1dT_{17}$ and cooled to 75°C. Five μ L of 10 x PCR buffer (Stratagene), 200 μ M dNTPs and 1.25 units of Taq polymerase were added, the mixture heated at 45°C for 2 min and 72°C for 40 min followed by 35 cycles of 94°C for 45 sec, 50°C for 25 sec, 72°C for 1.5 min and a final incubation at 72°C for 10 min. PCR products were separated by electrophoresis on 1% low melting agarose gels and the smear covering the range 600-800 bp fragments was excised and used in a second PCR amplification with 25 pmol of R_1 and POTSBE25 primers in a 50 μ L reaction (28 cycles of 94°C for 1 min, 50°C 1 min, 72°C 2 min). Products were purified by chloroform extraction and cloned into pT7 Blue. PCR was used to screen the colonies and the longest clones were sequenced.

The first round of RACE only extended the length of the SBE sequence approximately 100 bases, therefore a new A-tailed cDNA library was constructed using the class A SBE specific oligo POTSBE24 (10 pmol) in an attempt to recover longer RACE products. The first and second round PCR reactions were performed using new class A SBE primers (POTSBE 28 and 29 respectively) derived from the new sequence data. Conditions were as before except that the elongation step in the first PCR was for 3 min and the second PCR consisted of 28 cycles at 94 °C for 45 seconds, 55 °C for 25 sec and 72 °C for 1 min 45 sec.

Clones ranging in size from 400 bp to 1.4 kb were isolated and sequenced. The combined sequence of the longest RACE products and cDNA clones predicted a full length gene of about 3150 nucleotides, excluding the poly(A) tail (psbe 2con.seq in Fig. 8).

As the sequence of the 5' half of the gene was compiled from the sequence of several

RACE products generated using Taq polymerase, it was possible that the compiled sequence did not represent that of a single mRNA species and/or had nucleotide sequence changes. The 5' 1600 bases of the gene was therefore re-isolated by PCR using Ulma, a thermostable DNA polymerase which, because it possesses a 3'-5' exonuclease activity, has a lower error rate compared to Taq polymerase. Several PCR products were cloned and restriction mapped and found to differ in the number of *Hind* III, *Ssp* I, and *Eco*R I sites. These differences do not represent PCR artefacts as they were observed in clones obtained from independent PCR reactions (data not shown) and indicate that there are several forms of the class A SBE gene transcribed in potato tubers.

In order to ensure that the sequence of the full length cDNA clone was derived from a single mRNA species it was therefore necessary to PCR the entire gene in one piece. cDNA was prepared according to the RACE protocol except that the adaptor oligo $R_0R_1dT_{17}$ (5 pmol) was used as a primer and after synthesis the reaction was diluted to 200 μ L with TE pH 8 and stored at 4°C. Two μ L of the cDNA was used in a PCR reaction of 50 μ L using 25 pmol of class A SBE specific primers PBER1 and PBERT (see below), and thirty cycles of 94° for 1 min, 60°C for 1 min and 72°C for 3 min. If Taq polymerase was used the PCR products were cloned into pT7Blue whereas if Ulma polymerase was used the PCR products were purified by chloroform extraction, ethanol precipitation and kinased in a volume of 20 μ L (and then cloned into pBSSK IIP which had been cut with *Eco*RV and dephosphorylated). At least four classes of cDNA were isolated, which again differed in the presence or absence of *Hind* III, *Ssp* I and *Eco*R I sites. Three of these clones were sequenced fully, however one clone could not be isolated in sufficient quantity to sequence.

The sequence of one of the clones (number 19) is shown in Figure 5. The first methionine (initiation) codon starts a short open reading frame (ORF) of 7 amino acids which is out of frame with the next predicted ORF of 882 amino acids which has a molecular mass (Mr) of approximately 100 Kd. Nucleotides 6-2996 correspond to SBE sequence - the rest of the sequence shown is vector derived. Figure 6 shows a comparison of the most highly conserved part of the amino acid sequence of potato class A SBE (residues 180-871, top, row) and potato class B SBE (bottom row, residues 98-792); the middle row indicates the

degree of similarity, identical residues being denoted by the common letter, conservative changes by two dots and neutral changes by a single dot. Dashes indicate gaps introduced to optimise the alignment. The class A SBE protein has 44% identity over the entire length with potato class B SBE, and 56% identity therewith in the central conserved domain (Figure 6), as judged by the "Megalign" program (DNASTAR). However, Figure 7 shows a comparison between potato class A SBE (top row, residues 1-873) and pea class A SBE (bottom row, residues 1-861), from which it can be observed that cloned potato gene is more homologous to the class A pea enzyme, where the identity is 70 % over nearly the entire length, and this increases to 83 % over the central conserved region (starting at IPPP at position ~170). It is clear from this analysis that this cloned potato SBE gene belongs to the class A family of SBE genes.

An *E. coli* culture, containing the plasmid pSJ78 (which directs the expression of a full length potato SBE Class A gene), has been deposited (on 3rd January 1996) under the terms of the Budapest Treaty at The National Collections of Industrial and Marine Bacteria Limited (23 St Machar Drive, Aberdeen, AB2 1RY, United Kingdom), under accession number NCIMB 40781. Plasmid pSJ78 is equivalent to clone 19 described above. It represents a full length SBE A cDNA blunt-end ligated into the vector pBSSKIIIP.

Polymorphism of class A SBE genes

Sequence analysis of the other two full length class A SBE genes showed that they contain frameshift mutations and are therefore unable to encode full length proteins and indeed they were unable to complement the branching enzyme deficiency in the KV832 mutant (described below). An alignment of the full length DNA sequences is shown in Figure 8: "10con.seq" (Seq ID No. 12), "19con.seq" (Seq ID No. 14) and "11con.seq" (Seq ID No. 13) represent the sequence of full length clones 10, 19 and 11 obtained by PCR using the PBER1 and PBERT primers (see below), whilst "psbe2con.seq" (Seq ID No. 18) represents the consensus sequence of the RACE clones and cDNA clone 3.2.1. Those nucleotides which differ from the overall consensus sequence (not shown) are shaded. Dashes indicate gaps introduced to optimise the alignment. Apart from the frameshift mutations these clones are highly homologous. It should be noted that the 5' sequence of psbe2con is longer because this is the longest RACE product and it also contains several

changes compared to the other clones. The upstream methionine codon is still present in this clone but the upstream ORF is shortened to just 3 amino acids and in addition there is a 10 base deletion in the 5' untranslated leader.

The other significant area of variation is in the carboxy terminal region of the protein coding region. Closer examination of this area reveals a GAA trinucleotide repeat structure which varies in length between the four clones. These are typical characteristics of a microsatellite repeat region. The most divergent clone is #11 which has only one GAA triplet whereas clone 19 has eleven perfect repeats and the other two clones have five and seven GAA repeats. All of these deletions maintain the ORF but change the number of glutamic acid residues at the carboxy terminus of the protein.

Most of the other differences between the clones are single base changes. It is quite possible that some of these are PCR errors. To address this question direct sequencing of PCR fragments amplified from first strand cDNA was performed. Figure 9 shows the DNA sequence, and predicted amino acid sequence, obtained by such direct sequencing. Certain restriction sites are also marked. Nucleotides which could not be unambiguously assigned are indicated using standard IUPAC notation and, where this uncertainty affects the predicted amino acid sequence, a question mark is used. Sequence at the extreme 5' and 3' ends of the gene could not be determined because of the heterogeneity observed in the different cloned genes in these regions (see previous paragraph). However this can be taken as direct evidence that these differences are real and are not PCR or cloning artefacts.

There is absolutely no evidence for the frameshift mutations in the PCR derived sequence and it would appear that these mutations are an artefact of the cloning process, resulting from negative selection pressure in *E. coli*. This is supported by the fact that it proved extremely difficult to clone the full length PCR products intact as many large deletions were seen and the full length clones obtained were all cloned in one orientation (away from the LacZ promoter), perhaps suggesting that expression of the gene is toxic to the cells. Difficulties of this nature may have been responsible, at least in part, for the previous failure of other researchers to obtain the present invention.

A comparison of all the full length sequences is shown in Figure 10. In addition to clones 10, 11 and 19 are shown the sequences of a *Bgl* II - *Xho* I product cloned directly into the QE32 expression vector ("86CON.SEQ", Seq ID No. 16) and the consensus sequence of the directly sequenced PCR products ("pcrsbe2con.seq", Seq ID No. 17). Those nucleotides which differ from the consensus sequence (not shown) are shaded. Dashes indicate gaps introduced to optimise the alignment. There are 11 nucleotide differences predicted to be present in the mRNA population, which are indicated by asterisks above and below the sequence. The other differences are probably PCR artefacts or possibly sequencing errors.

Complementation of a branching enzyme deficient *E. coli* mutant

To determine if the isolated SBE gene encodes an active protein i.e. one that has branching enzyme activity, a complementation test was performed in the *E. coli* strain KV832. This strain is unable to make bacterial glycogen as the gene for the glycogen branching enzyme has been deleted (Keil *et al.*, 1987 Mol. Gen. Genet. 207, 294-301). When wild type cells are grown in the presence of glucose they synthesise glycogen (a highly branched glucose polymer) which stains a brown colour with iodine, whereas the KV832 cells make only a linear chain glucose polymer which stains blueish green with iodine. To determine if the cloned SBE gene could restore the ability of the KV832 cells to make a branched polymer, the clone pSJ90 (Seq ID No. 19) was used and constructed as below. The construct is a PCR-derived, substantially full length fragment (made using primers PBE 2B and PBE 2X, detailed below), which was cut with *Bgl* II and *Xho* I and cloned into the *Bam*H I / *Sal* I sites of the His-tag expression vector pQE32 (Qiagen). This clone, pSJ86, was sequenced and found to have a frameshift mutation of two bases in the 5' half of the gene. This frameshift was removed by digestion with *Nsi* I and *Sna*B I and replaced with the corresponding fragment from a Taq-generated PCR clone to produce the plasmid pSJ90 (sequence shown in Figure 12; the first 10 amino acids are derived from the expression vector). The polypeptide encoded by pSJ90 would be predicted to correspond to amino acids 46-882 of the full SBE coding sequence. The construct pSJ90 was transformed into the branching enzyme deficient KV832 cells and transformants were grown on solid PYG medium (0.85% KH_2PO_4 , 1.1% K_2HPO_4 , 0.6% yeast extract) containing 1.0% glucose. To test for complementation, a loop of cells was

scraped off and resuspended in 150 μ l of water, to which was added 15 μ l Lugol's solution (2g KI and 1g I₂ per 300ml water). It was found that the potato SBE fragment-transformed KV832 cells now stained a yellow-brown colour with iodine whereas control cells containing only the pQE32 vector continued to stain blue-green.

Expression of potato class A SBE in *E. coli*

Single colonies of KV832, containing one of the plasmids pQE32, pAGCR1 or pSJ90, were picked into 50ml of 2xYT medium containing carbenicillin, kanamycin and streptomycin as appropriate (100, 50 and 25 mg/L, respectively) in a 250ml flask and grown for 5 hours, with shaking, at 37°C. IPTG was then added to a final concentration of 1mM to induce expression and the flasks were further incubated overnight at 25°C. The cells were harvested by centrifugation and resuspended in 50 mM sodium phosphate buffer (pH 8.0), containing 300mM NaCl, 1mg/ml lysozyme and 1mM PMSF and left on ice for 1 hour. The cell lysates were then sonicated (3 pulses of 10 seconds at 40% power using a microprobe) and cleared by centrifugation at 12,000g for 10 minutes at 4°C. Cleared lysates were concentrated approximately 10 fold in a Centricon™ 30 filtration unit. Duplicate 10 μ l samples of the resulting extract were assayed for SBE activity by the phosphorylation stimulation method, as described in International Patent Application No. PCT/GB95/00634. In brief, the standard assay reaction mixture (0.2ml) was 200mM 2-(N-morpholino) ethanesulphonic acid (MES) buffer pH6.5, containing 100nCi of ¹⁴C glucose-1-phosphate at 50mM, 0.05 mg rabbit phosphorylase A, and *E. coli* lysate. The reaction mixture was incubated for 60 minutes at 30°C and the reaction terminated and glucan polymer precipitated by the addition of 1ml of 75% (v/v) methanol, 1% (w/v) potassium hydroxide, and then 0.1ml glycogen (10mg/ml). The results are presented below:

Construct	SBE Activity (cpm)
pQE32 (control)	1,829
pSJ90 (potato class A SBE)	14,327
pAGCR1 (pea class A SBE)	29,707

The potato class A SBE activity is 7-8 fold above background levels. It was concluded therefore that the potato class A SBE gene was able to complement the BE mutation in the

phosphorylation stimulation assay and that the cloned gene does indeed code for a protein with branching enzyme activity.

Oligonucleotides

The following synthetic oligonucleotides (Seq ID No.s 1-11 respectively) were used:

R ₀ R ₁ dT ₁₇	AAGGATCCGTCGACATCGATAATACGACTCACTATAGGGA(T) ₁₇
R ₀	AAGGATCCGTCGACATC
R ₁	GACATCGATAATACGAC
POTSBE24	CATCCAACCACCATCTCGCA
POTSBE25	TTGAGAGAAGATACCTAAGT
POTSBE28	ATGTTCAAGTCCATCTAAAGT
POTSBE29	AGAACAACAATTCCTAGCTC
PBER 1	GGGGCCTTGAACTCAGCAAT
PBERT	CGTCCCAGCATTCGACATAA
PBE 2B	CTTGGATCCTTGAACTCAGCAATTTG
PBE 2X	TAACTCGAGCAACGCGATCACAAGTTCGT

Example 2

Production of Transgenic Plants

Construction of plant transformation vectors with antisense starch branching enzyme genes

A 1200 bp *Sac* I - *Xho* I fragment, encoding approximately the -COOH half of the potato class A SBE (isolated from the rescued λ Zap clone 3.2.1), was cloned into the *Sac* I - *Sal* I sites of the plant transformation vector pSJ29 to create plasmid pSJ64, which is illustrated schematically in Figure 11. In the figure, the black line represents the DNA sequence. The broken line represents the bacterial plasmid backbone (containing the origin of replication and bacterial selection marker), which is not shown in full. The filled triangles on the line denote the T-DNA borders (RB = right border, LB = left border). Relevant restriction sites are shown above the black line, with the approximate distances (in kilobases) between the sites (marked by an asterisk) given by the numerals below the

line. The thinnest arrows indicate polyadenylation signals (pAnos = nopaline synthase, pAg7 = Agrobacterium gene 7), the arrows intermediate in thickness denote protein coding regions (SBE II = potato class A SBE, HYG = hygromycin resistance gene) and the thickest arrows represent promoter regions (P-2x35 = double CaMV 35S promoter, Pnos = nopaline synthase promoter). Thus pSJ64 contained the class A SBE gene fragment in an antisense orientation between the 2X 35S CaMV promoter and the nopaline synthase polyadenylation signal.

For information, pSJ29 is a derivative of the binary vector pGPTV-HYG (Becker *et al.*, 1992 Plant Molecular Biology 20, 1195-1197) modified as follows: an approximately 750 bp (*Sac* I, T4 DNA polymerase blunted - *Sal* I) fragment of pJIT60 (Guerineau *et al.*, 1992 Plant Mol. Biol. 18, 815-818) containing the duplicated cauliflower mosaic virus (CaMV) 35S promoter (Cabb-JI strain, equivalent to nucleotides 7040 to 7376 duplicated upstream of 7040 to 7433, Frank *et al.*, 1980 Cell 21, 285-294) was cloned into the *Hind* III (Klenow polymerase repaired) - *Sal* I sites of pGPTV-HYG to create pSJ29.

Plant transformation

Transformation was conducted on two types of potato plant explants; either wild type untransformed minitubers (in order to give single transformants containing the class A antisense construct alone) or minitubers from three tissue culture lines (which gave rise to plants #12, #15, #17 and #18 indicated in Table 1) which had already been successfully transformed with the class B (SBE I) antisense construct containing the tandem 35S promoter (so as to obtain double transformant plants, containing antisense sequences for both the class A and class B enzymes).

Details of the method of Agrobacterium transformation, and of the growth of transformed plants, are described in International Patent Application No. WO 95/26407, except that the medium used contained 3% sucrose (not 1%) until the final transfer and that the initial incubation with Agrobacterium (strain 3850) was performed in darkness. Transformants containing the class A antisense sequence were selected by growth in medium containing 15mg/L hygromycin (the class A antisense construct comprising the HYG gene, i.e. hygromycin phosphotransferase).

Transformation was confirmed in all cases by production of a DNA fragment from the antisense gene after PCR in the presence of appropriate primers and a crude extract of genomic DNA from each regenerated shoot.

Characterisation of starch from potato plants

Starch was extracted from plants as follows: potato tubers were homogenised in water for 2 minutes in a Waring blender operating at high speed. The homogenate was washed and filtered (initially through 2mm, then through 1mm filters) using about 4 litres of water per 100gms of tubers (6 extractions). Washed starch granules were finally extracted with acetone and air dried.

Starch extracted from singly transformed potato plants (class A/SBE II antisense, or class B/SBE I antisense), or from double transformants (class A/SBE II and class B/SBE I antisense), or from untransformed control plants, was partially characterised. The results are shown in Table 1. The table shows the amount of SBE activity (units/gram tissue) in tubers from each transformed plant. The endotherm peak temperature (°C) of starch extracted from several plants was determined by DSC, and the onset temperature (°C) of pasting was determined by reference to a viscoamylograph ("RVA"), as described in WO 95/26407. The viscoamylograph profile was as follows: step 1 - 50°C for 2 minutes; step 2 - increase in temperature from 50°C to 95°C at a rate of 1.5°C per minute; step 3 - holding at 95°C for 15 minutes; step 4 - cooling from 95°C to 50°C at a rate of 1.5°C per minute; and finally, step 5 - holding at 50°C for 15 minutes. Table 1 shows the peak, pasting and set-back viscosities in stirring number units (SNU), which is a measure of the amount of torque required to stir the suspensions. Peak viscosity may be defined for present purposes as the maximum viscosity attained during the heating phase (step 2) or the holding phase (step 3) of the viscoamylograph. Pasting viscosity may be defined as the viscosity attained by the starch suspensions at the end of the holding phase (step 3) of the viscoamylograph. Set-back viscosity may be defined as the viscosity of the starch suspension at the end of step 5 of the viscoamylograph.

A determination of apparent amylose content (% w/w) was also performed, using the iodometric assay method of Morrison & Laignelet (1983 *J. Cereal Sci.* 1, 9-20). The

results (percentage apparent amylose) are shown in Table 1. The untransformed and transformed control plants gave rise to starches having apparent amylose contents in the range 29(+/-3)%.

Generally similar values for amylose content were obtained for starch extracted from most of the singly transformed plants containing the class A (SBE II) antisense sequence. However, some plants (#152, 249) gave rise to starch having an apparent amylose content of 37-38%, notably higher than the control value. Starch extracted from these plants had markedly elevated pasting onset temperatures, and starch from plant 152 also exhibited an elevated endotherm peak temperature (starch from plant 249 was not tested by DSC).

Table 1

Sample description	Sample number	Tuber SBE activity (U/g starch)	DSC	Viscometry/ograph			(RVA)		Apparent amylose content (% w/w)	Phosphorus content (mg/100g)
				Peak temperature (°C)	Onset temperature (°C)	Peak viscosity (SNU)	Pasting viscosity (SNU)	Set-back viscosity (SNU)		
Untransformed control	146	7.6	65.6	65.5	545	161	260	31.2	66	
	243	22.2	nd	62.6	761	135	241	29.1		
AS-Class A SBE	152	12.7	69.5	70.9	467	380	529	37.5	69	
	249	13.9	nd	70.0	487	434	516	38.5		
AS-Class B SBE (17) (control)	145	0.7	66.9	66.6	669	177	305	29.8	111	
AS-Class B SBE (17) + AS-Class A SBE	150	0.6	74.0	68.0	214	214	303	53.1	196	
	161	0.5	73.0	76.6	349	324	616	40.9	208	
AS-Class B SBE (18) (control)	144	1.6	64.5	64.7	714	154	256	29.0	97	
AS-Class B SBE (18) + AS-Class A SBE	149	3.0	66.5	66.9	474	267	482	35.6	127	
	172	0.22	nd	65.4	707	167	290	26.6	130	
AS-Class B SBE (19) + AS-Class A SBE	201	0.10	nd	>85	no peak	12	13	66.4	210	
	206a	0.10	nd	>85	no peak	15	17	64.1	240	
	208	0.30	72.8-80.5	>85	no peak	14	19	62.8		
	202	0.02	nd	69.4	no peak	172	245	57.9		
	212	1.40	nd	78.0	308	286	541	49.5		
	220	1.40	nd	75.8	355	345	593	44.1		
AS-Class B SBE (13) (control)	170	0.2	nd	66.5	766	202	303	27.6		
AS-Class B SBE (12) + AS-Class A SBE	236	0.7	nd	85.0	no peak	23	14	60.4		
	236a	0.9	nd	91.2	no peak	139	192	56.7		
	230a	0.6	nd	77.6	244	239	450	46.2		

RVA profile
Pasting viscosity (47 min)
Set back viscosity (92 min)
SBE
SNU
nd

50°C (2 min), 50-85°C (1.5°C/min), 85°C (15 min), 85-60°C (1.5°C/min), 50°C (15 min)
at end of 50°C (2 min), 50-85°C (1.5°C/min), 85°C (15 min)
at end of profile
Starch Branching Enzyme
Instrument "Stirling Number Units" (arbitrary units)
not determined

29/1

Table 1

Sample description	Sample. number	Tuber SBE activity (U/g starch)	DSC	
			Peak temperature (°C)	Onset temperature (°C)
Untransformed control	146	7.6	65.8	65.5
	243	22.2	nd	62.6
AS-Class A SBE	152	12.7	69.5	70.9
	249	13.9	nd	70.0
AS-Class B SBE (17) (control)	145	0.7	66.9	66.8
AS-Class B SBE (17) + AS-Class A SBE	150	0.6	74.0	86.0
	161	0.5	73.0	76.6
AS-Class B SBE (18) (control)	144	1.6	64.5	64.7
AS-Class B SBE (18) + AS-Class A SBE	149	3.0	68.5	69.9

29/2

Viscoamylograph			Apparent amylose content (% w/w)	Phosphorus content (mg/100g)
Peak viscosity (SNU)	Pasting viscosity (SNU)	Set-back viscosity (SNU)		
545	161	280	31.2	68
761	135	241	29.1	
467	380	529	37.5	89
497	434	518	38.5	
669	177	305	29.8	111
214	214	303	53.1	198
349	324	618	40.9	206
714	154	258	29.0	97
474	267	482	35.6	127

AS-Class B SBE (15) (control)	172	0.22	nd	65.4
AS-Class B SBE (15) + AS-Class A SBE	201	0.10	nd	>95
	208a	0.10	nd	>95
	208	0.30	72.8-80.5	>95
	202	0.02	nd	89.4
	212	1.40	nd	78.0
	220	1.40	nd	75.8
AS-Class B SBE (12) (control)	170	0.2	nd	66.5
AS-Class B SBE (12) + AS-Class A SBE	236	0.7	nd	95.0
	236a	0.9	nd	91.2
	230a	0.8	nd	77.6

RVA profile

Pasting viscosity (47 min)

Set-back viscosity (92 min)

SBE

SNU

nd

50°C (2 min), 50-95°C (1.5°C/min), 95°C (15 min), 95-50°C (1.5°C/min), 50°C (15 min)

at end of 50°C (2min), 50-95°C (1.5°C/min), 95°C (15 min)

at end of profile

Starch Branching Enzyme

Instrument "Stirring Number Units" (arbitrary units)

not detected

29/4

707	167	290	28.8	130
no peak	12	13	66.4	210
no peak	15	17	64.1	
no peak	14	19	62.8	240
no peak	172	245	57.9	
308	296	541	49.5	
355	345	593	44.1	
768	202	303	27.8	
no peak	23	14	60.4	
no peak	139	192	56.7	
244	239	450	48.2	

SUBSTITUTE SHEET (RULE 26)

It should be noted that, even if other single transformants were not to provide starch with an altered amylose/amylopectin ratio, the starch from such plants might still have different properties relative to starch from conventional plants (e.g. different average molecular weight or different amylopectin branching patterns), which might be useful.

Double transformant plants, containing antisense sequences for both the class A and class B enzymes, had greatly reduced SBE activity (units/gm) compared to untransformed plants or single anti-sense class A transformants, (as shown in Table 1). Moreover, certain of the double transformant plants contained starch having very significantly altered properties. For example, starch extracted from plants #201, 202, 208, 208a, 236 and 236a had drastically altered amylose/amylopectin ratios, to the extent that amylose was the main constituent of starch from these plants. The pasting onset temperatures of starch from these plants were also the most greatly increased (by about 25-30°C). Starch from plants such as #150, 161, 212, 220 and 230a represented a range of intermediates, in that such starch displayed a more modest rise in both amylose content and pasting onset temperature. The results would tend to suggest that there is generally a correlation between % amylose content and pasting onset temperature, which is in agreement with the known behaviour of starches from other sources, notably maize.

The marked increase in amylose content obtained by inhibition of class A SBE alone, compared to inhibition of class B SBE alone (see PCT/GB95/00634) might suggest that it would be advantageous to transform plants first with a construct to suppress class A SBE expression (probably, in practice, an antisense construct), select those plants giving rise to starch with the most altered properties, and then to re-transform with a construct to suppress class B SBE expression (again, in practice, probably an antisense construct), so as to maximise the degree of starch modification.

In addition to pasting onset temperatures, other features of the viscoamylograph profile e.g. for starches from plants #149, 150, 152, 161, 201, 236 and 236a showed significant differences to starches from control plants, as illustrated in Figure 13. Referring to Figure 13, a number of viscoamylograph traces are shown. The legend is as follows: shaded box - normal potato starch control (29.8% amylose content); shaded circle - starch from plant

149 (35.6% amylose); shaded triangle, pointing upwards - plant 152 (37.5%); shaded triangle, pointing downwards - plant 161 (40.9%); shaded diamond - plant 150 (53.1%); unshaded box - plant 236a (56.7%); unshaded circle - plant 236 (60.4%); unshaded triangle, pointing upwards - plant 201 (66.4%); unshaded triangle, pointing downwards - Hylon V starch, from maize (44.9 % amylose). The thin line denotes the heating profile.

With increasing amylose content, peak viscosities during processing to 95°C decrease, and the drop in viscosity from the peak until the end of the holding period at 95°C also generally decreases (indeed, for some of the starch samples there is an increase in viscosity during this period). Both of these results are indicative of reduced granule fragmentation, and hence *increased* granule stability during pasting. This property has not previously been available in potato starch without extensive prior chemical or physical modification. For applications where a maximal viscosity after processing to 95°C is desirable (i.e. corresponding to the viscosity after 47 minutes in the viscoamylograph test), starch from plant #152 would be selected as starches with both lower (Controls, #149) and higher (#161, #150) amylose contents have lower viscosities following this gelatinisation and pasting regime (Figure 13 and Table 1). It is believed that the viscosity at this stage is determined by a combination of the extent of granule swelling and the resistance of swollen granules to mechanical fragmentation. For any desired viscosity behaviour, one skilled in the art would select a potato starch from a range containing different amylose contents produced according to the invention by performing suitable standard viscosity tests.

Upon cooling pastes from 95°C to 50°C, potato starches from most plants transformed in accordance with the invention showed an increase in viscoamylograph viscosity as expected for partial reassociation of amylose. Starches from plants #149, 152 and 161 all show viscosities at 50°C significantly in excess of those for starches from control plants (Figure 13 and Table 1). This contrasts with the effect of elevated amylose contents in starches from maize plants (Figure 2) which show very low viscosities throughout the viscoamylograph test. Of particular note is the fact that, for similar amylose contents, starch from potato plant 150 (53% amylose) shows markedly increased viscosity compared with Hylon 5 starch (44.9% amylose) as illustrated in Figure 13. This demonstrates that

useful properties which require elevated (35% or greater) amylose levels can be obtained by processing starches from potato plants below 100°C, whereas more energy-intensive processing is required in order to generate similarly useful properties from high amylose starches derived from maize plants.

Final viscosity in the viscoamylograph test (set-back viscosity after 92 minutes) is greatest for starch from plant #161 (40.9% amylose) amongst those tested (Figure 13 and Table 1). Decreasing final viscosities are obtained for starches from plant #152 (37.5% amylose), #149 (35.6% amylose) and #150 (53.1% amylose). Set-back viscosity occurs where amylose molecules, exuded from the starch granule during pasting, start to re-associate outside the granule and form a viscous gel-like substance. It is believed that the set-back viscosity values of starches from transgenic potato plants represent a balance between the inherent amylose content of the starches and the ability of the amylose fraction to be exuded from the granule during pasting and therefore be available for the reassociation process which results in viscosity increase. For starches with low amylose content, increasing the amylose content tends to make more amylose available for re-association, thus increasing the set-back viscosity. However, above a threshold value, increased amylose content is thought to inhibit granule swelling, thus preventing exudation of amylose from the starch granule and reducing the amount of amylose available for re-association. This is supported by the RVA results obtained for the very high amylose content potato starches seen in the viscoamylograph profiles in Figure 13. For any desired viscosity behaviour following set-back or retrogradation to any desired temperature over any desired timescale, one skilled in the art would select a potato starch from a range containing different amylose contents produced according to the invention by performing standard viscosity tests.

Further experiments with starch from plants #201 and 208 showed that this had an apparent amylose content of over 62% (see Table 1). Viscoamylograph studies showed that starch from these plants had radically altered properties and behaved in a manner similar to hylon 5 starch from maize plants (Figure 13). Under the conditions employed in the viscoamylograph, this starch exhibited extremely limited (nearly undetectable) granule swelling. Thus, for example, unlike starch from control plants, starch from plants

201, 208 and 208a did not display a clearly defined pasting viscosity peak during the heating phase. Microscopic analysis confirmed that the starch granule structure underwent only minor swelling during the experimental heating process. This property may well be particularly useful in certain applications, as will be apparent to those skilled in the art.

Some re-grown plants have so far been found to increase still further the apparent amylose content of starch extracted therefrom. Such increases may be due to:-

- i) Growth and development of the first generation transformed plants may have been affected to some degree by the exogenous growth hormones present in the tissue culture system, which exogenous hormones were not present during growth of the second generation plants; and
- ii) Subsequent generations were grown under field conditions, which may allow for attainment of greater maturity than growth under laboratory conditions, it being generally held that amylose content of potato starch increases with maturity of the potato tuber.

Accordingly, it should be possible to obtain potato plants giving rise to tubers with starch having an amylose content in excess of the 66% level so far attained, simply by analysing a greater number of transformed plants and/or by re-growing transgenic plants through one or more generations under field conditions.

Table 1 shows that another characteristic of starch which is affected by the presence of anti-sense sequences to SBE is the phosphorus content. Starch from untransformed control plants had a phosphorus content of about 60-70mg/100gram dry weight (as determined according to the AOAC Official Methods of Analysis, 15th Edition, Method 948.09 "Phosphorus in Flour"). Introduction into the plant of an anti-sense SBE B sequence was found to cause a modest increase (about two-fold) in phosphorus content, which is in agreement with the previous findings reported at scientific meetings. Similarly, anti-sense to SBE A alone causes only a small rise in phosphorus content relative to untransformed controls. However, use of anti-sense to both SBE A and B in combination results in up to a four-fold increase in phosphorus content, which is far greater than any *in planta* phosphorus content previously demonstrated for potato starch.

This is useful in that, for certain applications, starch must be phosphorylated *in vitro* by

chemical modification. The ability to obtain potato starch which, as extracted from the plant, already has a high phosphorus content will reduce the amount of *in vitro* phosphorylation required suitably to modify the starch. Thus, in another aspect the invention provides potato starch which, as extracted from the plant, has a phosphorus content in excess of 200mg/100gram dry weight starch. Typically the starch will have a phosphorus content in the range 200 - 240mg/100gram dry weight starch.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: National Starch and Chemical Investment Holding Corporation
- (B) STREET: 501 Silverside Road, Suite 27
- (C) CITY: Wilmington
- (D) STATE: Delaware
- (E) COUNTRY: United States of America
- (F) POSTAL CODE (ZIP): 19809

(ii) TITLE OF INVENTION: Improvements in or Relating to Plant Starch Composition

(iii) NUMBER OF SEQUENCES: 20

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 57 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

AAGGATCCGT CGACATCGAT AATACGACTC ACTATAGGGA TTTTTTTTTT TTTTTT

57

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

AAGGATCCGT CGACATC

17

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GACATCGATA ATACGAC

17

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CATCCAACCA CCATCTCGCA

20

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TTGAGAGAAG ATACCTAAGT

20

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

ATGTTCAAGT CATCTAAAGT

20

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

AGAACAACAA TTCCTAGCTC

20

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGGGCCTTGA ACTCAGCAAT

20

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CGTCCCAGCA TTCGACATAA

20

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CTTGGATCCT TGAAGTCAGC AATTTG

26

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

TAACTCGAGC AACGCGATCA CAAGTTCGT

29

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3003 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GATGGGGCCT TGAAGTCAGC AATTTGACAC TCAGTTAGTT AACTGCCAT CACTTATCAG	60
ATCTCTATTT TTTCTCTTAA TTCCAACCAA GGAATGAATA AAAAGATAGA TTTGTAAAAA	120
CCCTAAGGAG AGAAGAAGAA AGATGGTGTA TAACTCTCT GGAGTTCGTT TTCCTACTGT	180
TCCATCAGTG TACAAATCTA ATGGATTCAG CAGTAATGGT GATCGGAGGA ATGCTAATAT	240
TTCTGTATTC TTGAAAAAC ACTCTCTTTC ACGGAAGATC TTGGCTGAAA AGTCTTCTTA	300
CAATTCCGAA TCCCGACCTT CTACAATTGC AGCATCGGGG AAAGTCCTTG TGCCTGGAAT	360
CCAGAGTGAT AGCTCCTCAT CCTCAACAGA TCAATTTGAG TTCGCTGAGA CATCTCCAGA	420
AAATTCCCCA GCATCAACTG ATGTAGATAG TTCAACAATG GAACACGCTA GCCAGATTAA	480
AACTGAGAAC GATGACGTTG AGCCGTCAAG TGATCTTACA GGAAGTGTTG AAGAGCTGGA	540
TTTTGCTTCA TCACTACAAC TACAAGAAGG TGGTAACTG GAGGAGTCTA AAACATTAAA	600
TACTTCTGAA GAGACAATTA TTGATGAATC TGATAGGATC AGAGAGAGGG GCATCCCTCC	660
ACCTGGACTT GGTCAGAAGA TTTATGAAAT AGACCCCTT TTGACAACT ATCGTCAACA	720
CCTTGATTAC AGGTATTCAC AGTACAAGAA ACTGAGGGAG GCAATTGACA AGTATGAGGG	780
TGGTTTGGAA GCTTTTTCTC GTGGTTATGA AAGAATGGGT TCACTCGTA GTGCTACAGG	840
TATCACTTAC CGTGAGTGGG CTCCTGGTGC CCAGTCAGCT GCCCTCATTG GGGATTTCAA	900
CAATTGGGAC GCAAATGCTG ACTTTATGAC TCGGAATGAA TTTGGTGTCT GAGAGATTTT	960
TCTGCCAAAT AATGTGGATG GTTCTCCTGC AATTCCTCAT GGGTCCAGAG TGAAGATACG	1020
TATGGAACT CCATCAGGTG TTAAGGATTC CATTCTGCT TGGATCACT ACTCTTTACA	1080
GCTTCCTGAT GAAATTCCAT ATAATGGAAT ATATTATGAT CCACCCGAAG AGGAGAGGTA	1140
TATCTTCCAA CACCCACGGC CAAAGAAACC AAAGTCGGTG AGAATATATG AATCTCATAT	1200
TGGAATGAGT AGTCCGGAGC CTAAAATTAA CTCATACGTG AATTTTAGAG ATGAAGTTCT	1260
TCCTCGCATA AAAAAAGCTT GGGTACAATG CGGTGCAAAT TATGGCTATT CAAGAGCATT	1320
CTTATTATGC TAGTTTTGGT TATCATGTCA CAAATTTTTT TGCACCAAGC AGCCGTTTTG	1380

GAACGCCCCGA CGACCTTAAG TCTTTGATTG ATAAAGCTCA TGAGCTAGGA ATTGTTGTTC	1440
TCATGGACAT TGTTACAGC CATGCATCAA ATAATACTTT AGATGGACTG AACATGTTTG	1500
ACGGCACAGA TAGTTGTTAC TTTCACCTCTG GAGCTCGTGG TTATCATTGG ATGTGGGATT	1560
TCCGCCTCTT TAACTATGGA AACTGGGAGG TACTTAGGTA TCTTCTCTCA AATGCGAGAT	1620
GGTGGTTGGA TGAGTTCAAA TTTGATGGAT TTAGATTTGA TGGTGTGACA TCAATGATGT	1680
GTACTCACCA CGGATTATCG GTGGGATTCA CTGGGAAC TA CGAGGAATAC TTTGGACTCG	1740
CAACTGATGT GGATGCTGTT GTGTATCTGA TGCTGGTCAA CGATCTTATT CATGGGCTTT	1800
TCCCAGATGC AATTACCATT GGTGAAGATG TTAGCGGAAT GCCGACATTT TGTGTTCCCG	1860
TTCAAGATGG GGGTGTGGC TTTGACTATC GGCTGCATAT GGCAATTGCT GATAAATGGA	1920
TTGAGTTGCT CAAGAAACGG GATGAGGATT GGAGAGTGGG TGATATTGTT CATACACTGA	1980
CAAATAGAAG ATGGTCGGAA AAGTGTGTTT CATACGCTGA AAGTCATGAT CAAGCTCTAG	2040
TCGGTGATAA AACTATAGCA TTCTGGCTGA TGGACAAGGA TATGTATGAT TTTATGGCTC	2100
TGGATAGACC GTCAACATCA TTAATAGATC GTGGGATAGC ATTACACAAG ATGATTAGGC	2160
TTGTAAGTAT GGGATTAGGA GGAGAAGGGT ACCTAAATTT CATGGGAAAT GAATTCGGCC	2220
ACCCTGAGTG GATTGATTC CCTAGGGCTG AACAACACCT CTCTGATGGC TCAGTAATTC	2280
CCAGAAACCA ATTCAGTTAT GATAAATGCA GACGGAGATT TGACCTGGGA GATGCAGAAT	2340
ATTTAAGATA CCGTGGGTTG CAAGAATTTG ACCGGGCTAT GCAGTATCTT GAAGATAAAT	2400
ATGAGTTTAT GACTTCAGAA CACCAGTTCA TATCACGAAA GGATGAAGGA GATAGGATGA	2460
TTGTATTTGA AAAAGGAAAC CTAGTTTTTG TCTTTAATTT TCACTGGACA AAAGGCTATT	2520
CAGACTATCG CATAGGCTGC CTGAAGCCTG GAAAATACAA GGTTGCCTTG GACTCAGATG	2580
ATCCACTTTT TGGTGGCTTC GGGAGAATTG ATCATAATGC CGAATATTTT ACCTTTGAAG	2640
GATGGTATGA TGATCGTCCT CGTTCAATTA TGGTGTATGC ACCTAGTAGA ACAGCAGTGG	2700
TCTATGCACT AGTAGACAAA GAAGAAGAAG AAGAAGAAGA AGTAGCAGTA GTAGAAGAAG	2760
TAGTAGTAGA AGAAGAATGA ACGAACTTGT GATCGCGTTG AAAGATTTGA ACGCCACATA	2820
GAGCTTCTTG ACGTATCTGG CAATATTGCA TTAGTCTTGG CGGAATTTCA TGTGACAACA	2880
GGTTTGCAAT TCTTTCCACT ATTAGTAGTG CAACGATATA CGCAGAGATG AAGTGCTGAA	2940
CAAAAACATA TGTAATTCG ATGAATTTAT GTCGAATGCT GGGACGATCG AATTCCTGCA	3000
GCC	3003

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2975 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TTGATGGGCC TTGAACTCAG CAATTTGACA CTCAGTTAGT TAACTCCTA TCACTTATCA	60
GATCTCTATT TTTTCTCTTA ATTCCAACCA GGGGAATGAA TAAAAGGATA GATTTGTAAA	120
AACCCTAAGG AGAGAAGAAG AAAGATGGTG TATATACTCT CTGGAGTTCG TTTTCCTACT	180
GTTCCATCAG TGTACAAATC TAATGGATTG AGCAGTAATG GTGATCGGAG GAATGCTAAT	240
GTTTCTGTAT TCTTGAAAAA GCACTCTCTT TCACGGAAGA TCTTGGCTGA AAAGTCTTCT	300
TACAATTCCG AATTCCGACC TTCTACAGTT GCAGCATCGG GGAAAGTCCT TGTGCCTGGA	360
ACCCAGAGTG ATAGCTCCTC ATCCTCAACA GACCAATTTG AGTTCCTGA GACATCTCCA	420
GAAAAATCCC CAGCATCAAC TGATGTAGAT AGTTCAACAA TGGAACACGC TAGCCAGATT	480
AAAACTGAGA ACGATGACGT TGAGCCGTCA AGTGATCTTA CAGGAAGTGT TGAAGAGCTG	540
GATTTTGCTT CATCACTACA ACTACAAGAA GGTGGTAAAC TGGAGGAGTC TAAACATTA	600
AATACTTCTG AAGAGACAAT TATTGATGAA TCTGATAGGA TCAGAGAGAG GGGCATCCCT	660
CCACCTGGAC TTGGTCAGAA GATTTATGAA ATAGACCCCC TTTTGACAAA CTATCGTCAA	720
CACCTTGATT ACAGGTATTC ACAGTACAAG AAAGTGAGGG AGGCAATTGA CAAGTATGAG	780
GGTGGTTTGG AAGCTTTTCT CGTGGTTATG AAAAAATGGG TTTCACTCGT AGTGCTACAG	840
GTATCACTTA CCGTGAGTGG GCTCCTGGTG CCCAGTCAGC TGCCCTCATT GGAGATTTCA	900
ACAATTGGGA CGCAAATGCT GACATTATGA CTCGGAATGA ATTTGGTGTC TGGGAGATTT	960
TTCTGCCAAA TAATGTGGAT GGTTCCTCTG CAATTCCTCA TGGGTCCAGA GTGAAGATAC	1020
GTATGGACAC TCCATCAGGT GTTAAGGATT CCATTCCTGC TTGGATCAAC TACTCTTTAC	1080
AGCTTCCTGA TGAAATTCCA TATAATGGAA TATATTATGA TCCACCCGAA GAGGAGAGGT	1140
ATATCTTCCA ACACCCACGG CCAAAGAAAC CAAAGTCGCT GAGAATATAT GAATCTCATA	1200
TTGGAATGAG TAGTCCGGAG CCTAAAATTA ACTCATACGT GAATTTTAGA GATGAAGTTC	1260
TTCTCGCAT AAAAAAGCTT GGGTACAATG CGCTGCGAAT TATGGCTATT CAAGAGCATT	1320
CTTATTATGC TAGTTTTGGT TATCATGTCA CAAATTTTTT TGCACCAAGC AGCCGTTTTG	1380

GAACGCCCCGA CGACCTTAAG TCTTCGATTG ATAAAGCTCA TGAGCTAGGA ATTGTTGTTC 1440
TCATGGACAT CGTTCACAGC CATGCATCAA ATAATACTTT AGATGGACTG AACATGTTTG 1500
ACGGCACCGA TAGTTGTTAC TTTCACTCTG GAGCTCGTGG TTATCATTGG ATGTGGGATT 1560
CCGCTCTTTT AACTATGGAA ACTGGGAGGT ACTTAGGTAT CTTCTCTCAA ATGCGAGATG 1620
GTGGTTGGAT GAGTTCAAAT TTGATGGATT TAGATTCGAT GGTGTGACAT CAATGATGTA 1680
TACTCACCAC GGATTATCGG TGGGATTCAC TGGGAACTAC GAGGAATACT TTGGACTCGC 1740
AACTGATGTG GATGCTGTTG TGTATCTGAT GCTGGTCAAC GATCTTATTC ATAGGCTTTT 1800
CCCAGATGCA ATTACCATTG GTGAAGATGT TAGCGGAATG CCGACATTTT GTATTCCCGT 1860
TCAAGATGGG GGTGTTGGCT TTGACTATCG GCTGCATATG GCAATTGCTG ATAAATGGAT 1920
TGAGTTGCTC AAGAAACGGG ATGAGGATTG GAGAGTGGGT GATATTGTTC ATACACTGAC 1980
AAATAGAAGA TGGTCGGAAA AGTGTGTTTC ATACGCTGAA AGTCATGATC AAGCTCTAGT 2040
CGGTGATAAA ACTATAGCAT TCTGGCTGAT GGACAAGGAT ATGTATGATT TTATGGCTCT 2100
GGATAGACCG CCAACATCAT TAATAGATCG TGGGATAGCA TTGCACAAGA TGATTAGGCT 2160
TGTAACATATG GGATTAGGAG GAGAAGGGTA CCTAAATTTT ATGGGAAATG AATTCGGCCA 2220
CCCTGAGTGG ATTGATTTCC CTAGGGCTGA GCCACACCTT TCTGATGGCT CAGTAATTCC 2280
CGGAAACCAA TTCAGTTATG ATAAATGCAG ACGGAGATTT GACCTGGGAG ATGCAGAATA 2340
TTTAAGATAC CATGGGTTAC AAGAATTTGA CTGGGCTATG CAGTATCTTG AAGATAAATA 2400
TGAGTTTATG ACTTCAGAAC ACCAGTTCAT ATCACGAAAG GATGAAGGAG ATAGGATGAT 2460
TGTATTTGAA AGAGGAAACC TAGTTTTCTG CTTTAATTTT CACTGGACAA ATAGCTATTC 2520
AGACTATCGC ATAGGCTGCC TGAAGCCTGG AAAATACAAG GTTGTCTTGG ACTCAGATGA 2580
TCCACTTTTT GGTGGCTTCG GGAGAATTGA TCATAATGCC GAATATTTCA CCTCTGAAGG 2640
ATCGTATGAT GATCGTCCTT GTTCAATTAT GGTGTATGCA CCTAGTAGAA CAGCAGTGGT 2700
CTATGCACTA GTAGACAAAC TAGAAGTAGC AGTAGTAGAA GAACCCATTG AAGAATGAAC 2760
GAACTTGTGA TCGCGTTGAA AGATTTGAAC GTTACTTGGT CATCCACATA GAGCTTCTTG 2820
ACATCAGTCT TGGCGGAATT GCATGTGACA ACAAGGTTTG CAGTTCTTTC CACTATTAGT 2880
AGTCCACCGA TATACGCAGA GATGAAGTGC TGAACAAACA TATGTAAAAT CGATGAATTT 2940
ATGTCGAATG CTGGGACGAT CGAATTCCTG CAGCC 2975

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3033 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 145..2790

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TTGATGGGGC CTTGAACTCA GCAATTTGAC ACTCAGTTAG TTACACTCCT ATCACTTATC	60
AGATCTCTAT TTTTCTCTT AATTCCAACC AAGGAATGAA TAAAAGGATA GATTTGTAAA	120
AACCCTAAGG AGAGAAGAAG AAAG ATG GTG TAT ACA CTC TCT GGA GTT CGT	171
Met Val Tyr Thr Leu Ser Gly Val Arg	
1 5	
TTT CCT ACT GTT CCA TCA GTG TAC AAA TCT AAT GGA TTC AGC AGT AAT	219
Phe Pro Thr Val Pro Ser Val Tyr Lys Ser Asn Gly Phe Ser Ser Asn	
10 15 20 25	
GGT GAT CGG AGG AAT GCT AAT GTT TCT GTA TTC TTG AAA AAG CAC TCT	267
Gly Asp Arg Arg Asn Ala Asn Val Ser Val Phe Leu Lys Lys His Ser	
30 35 40	
CTT TCA CGG AAG ATC TTG GCT GAA AAG TCT TCT TAC AAT TCC GAA TTC	315
Leu Ser Arg Lys Ile Leu Ala Glu Lys Ser Ser Tyr Asn Ser Glu Phe	
45 50 55	
CGA CCT TCT ACA GTT GCA GCA TCG GGG AAA GTC CTT GTG CCT GGA ACC	363
Arg Pro Ser Thr Val Ala Ala Ser Gly Lys Val Leu Val Pro Gly Thr	
60 65 70	
CAG AGT GAT AGC TCC TCA TCC TCA ACA GAC CAA TTT GAG TTC ACT GAG	411
Gln Ser Asp Ser Ser Ser Ser Ser Thr Asp Gln Phe Glu Phe Thr Glu	
75 80 85	
ACA TCT CCA GAA AAT TCC CCA GCA TCA ACT GAT GTA GAT AGT TCA ACA	459
Thr Ser Pro Glu Asn Ser Pro Ala Ser Thr Asp Val Asp Ser Ser Thr	
90 95 100 105	
ATG GAA CAC GCT AGC CAG ATT AAA ACT GAG AAC GAT GAC GTT GAG CCG	507
Met Glu His Ala Ser Gln Ile Lys Thr Glu Asn Asp Asp Val Glu Pro	
110 115 120	
TCA AGT GAT CTT ACA GGA AGT GTT GAA GAG CTG GAT TTT GCT TCA TCA	555
Ser Ser Asp Leu Thr Gly Ser Val Glu Glu Leu Asp Phe Ala Ser Ser	
125 130 135	

CTA	CAA	CTA	CAA	GAA	GGT	GGT	AAA	CTG	GAG	GAG	TCT	AAA	ACA	TTA	AAT	603
Leu	Gln	Leu	Gln	Glu	Gly	Gly	Lys	Leu	Glu	Glu	Ser	Lys	Thr	Leu	Asn	
		140					145					150				
ACT	TCT	GAA	GAG	ACA	ATT	ATT	GAT	GAA	TCT	GAT	AGG	ATC	AGA	GAG	AGG	651
Thr	Ser	Glu	Glu	Thr	Ile	Ile	Asp	Glu	Ser	Asp	Arg	Ile	Arg	Glu	Arg	
		155				160					165					
GGC	ATC	CCT	CCA	CCT	GGA	CTT	GGT	CAG	AAG	ATT	TAT	GAA	ATA	GAC	CCC	699
Gly	Ile	Pro	Pro	Pro	Gly	Leu	Gly	Gln	Lys	Ile	Tyr	Glu	Ile	Asp	Pro	
					175					180					185	
CTT	TTG	ACA	AAC	TAT	CGT	CAA	CAC	CTT	GAT	TAC	AGG	TAT	TCA	CAG	TAC	747
Leu	Leu	Thr	Asn	Tyr	Arg	Gln	His	Leu	Asp	Tyr	Arg	Tyr	Ser	Gln	Tyr	
				190					195					200		
AAG	AAA	CTG	AGG	GAG	GCA	ATT	GAC	AAG	TAT	GAG	GGT	GGT	TTG	GAA	GCC	795
Lys	Lys	Leu	Arg	Glu	Ala	Ile	Asp	Lys	Tyr	Glu	Gly	Gly	Leu	Glu	Ala	
			205					210					215			
TTT	TCT	CGT	GGT	TAT	GAA	AAA	ATG	GGT	TTC	ACT	CGT	AGT	GCT	ACA	GGT	843
Phe	Ser	Arg	Gly	Tyr	Glu	Lys	Met	Gly	Phe	Thr	Arg	Ser	Ala	Thr	Gly	
		220					225					230				
ATC	ACT	TAC	CGT	GAG	TGG	GCT	CTT	GGT	GCC	CAG	TCA	GCT	GCC	CTC	ATT	891
Ile	Thr	Tyr	Arg	Glu	Trp	Ala	Leu	Gly	Ala	Gln	Ser	Ala	Ala	Leu	Ile	
		235				240				245						
GGA	GAT	TTC	AAC	AAT	TGG	GAC	GCA	AAT	GCT	GAC	ATT	ATG	ACT	CGG	AAT	939
Gly	Asp	Phe	Asn	Asn	Trp	Asp	Ala	Asn	Ala	Asp	Ile	Met	Thr	Arg	Asn	
					255					260					265	
GAA	TTT	GGT	GTC	TGG	GAG	ATT	TTT	CTG	CCA	AAT	AAT	GTG	GAT	GGT	TCT	987
Glu	Phe	Gly	Val	Trp	Glu	Ile	Phe	Leu	Pro	Asn	Asn	Val	Asp	Gly	Ser	
				270					275					280		
CCT	GCA	ATT	CCT	CAT	GGG	TCC	AGA	GTG	AAG	ATA	CGT	ATG	GAC	ACT	CCA	1035
Pro	Ala	Ile	Pro	His	Gly	Ser	Arg	Val	Lys	Ile	Arg	Met	Asp	Thr	Pro	
			285					290					295			
TCA	GGT	GTT	AAG	GAT	TCC	ATT	CCT	GCT	TGG	ATC	AAC	TAC	TCT	TTA	CAG	1083
Ser	Gly	Val	Lys	Asp	Ser	Ile	Pro	Ala	Trp	Ile	Asn	Tyr	Ser	Leu	Gln	
		300					305					310				
CTT	CCT	GAT	GAA	ATT	CCA	TAT	AAT	GGA	ATA	CAT	TAT	GAT	CCA	CCC	GAA	1131
Leu	Pro	Asp	Glu	Ile	Pro	Tyr	Asn	Gly	Ile	His	Tyr	Asp	Pro	Pro	Glu	
		315				320					325					
GAG	GAG	AGG	TAT	ATC	TTC	CAA	CAC	CCA	CGG	CCA	AAG	AAA	CCA	AAG	TCG	1179
Glu	Glu	Arg	Tyr	Ile	Phe	Gln	His	Pro	Arg	Pro	Lys	Lys	Pro	Lys	Ser	
					335					340					345	
CTG	AGA	ATA	TAT	GAA	TCT	CAT	ATT	GGA	ATG	AGT	AGT	CCG	GAG	CCT	AAA	1227
Leu	Arg	Ile	Tyr	Glu	Ser	His	Ile	Gly	Met	Ser	Ser	Pro	Glu	Pro	Lys	
				350					355					360		

ATT AAC TCA TAC GTG AAT TTT AGA GAT GAA GTT CTT CCT CGC ATA AAA	1275
Ile Asn Ser Tyr Val Asn Phe Arg Asp Glu Val Leu Pro Arg Ile Lys	
365 370 375	
AAG CTT GGG TAC AAT GCG CTG CAA ATT ATG GCT ATT CAA GAG CAT TCT	1323
Lys Leu Gly Tyr Asn Ala Leu Gln Ile Met Ala Ile Gln Glu His Ser	
380 385 390	
TAT TAC GCT AGT TTT GGT TAT CAT GTC ACA AAT TTT TTT GCA CCA AGC	1371
Tyr Tyr Ala Ser Phe Gly Tyr His Val Thr Asn Phe Phe Ala Pro Ser	
395 400 405	
AGC CGT TTT GGA ACG CCC GAC GAC CTT AAG TCT TTG ATT GAT AAA GCT	1419
Ser Arg Phe Gly Thr Pro Asp Asp Leu Lys Ser Leu Ile Asp Lys Ala	
410 415 420 425	
CAT GAG CTA GGA ATT GTT GTT CTC ATG GAC ATT GTT CAC AGC CAT GCA	1467
His Glu Leu Gly Ile Val Val Leu Met Asp Ile Val His Ser His Ala	
430 435 440	
TCA AAT AAT ACT TTA GAT GGA CTG AAC ATG TTT GAC TGC ACC GAT AGT	1515
Ser Asn Asn Thr Leu Asp Gly Leu Asn Met Phe Asp Cys Thr Asp Ser	
445 450 455	
TGT TAC TTT CAC TCT GGA GCT CGT GGT TAT CAT TGG ATG TGG GAT TCC	1563
Cys Tyr Phe His Ser Gly Ala Arg Gly Tyr His Trp Met Trp Asp Ser	
460 465 470	
CGC CTC TTT AAC TAT GGA AAC TGG GAG GTA CTT AGG TAT CTT CTC TCA	1611
Arg Leu Phe Asn Tyr Gly Asn Trp Glu Val Leu Arg Tyr Leu Leu Ser	
475 480 485	
AAT GCG AGA TGG TGG TTG GAT GCG TTC AAA TTT GAT GGA TTT AGA TTT	1659
Asn Ala Arg Trp Trp Leu Asp Ala Phe Lys Phe Asp Gly Phe Arg Phe	
490 495 500 505	
GAT GGT GTG ACA TCA ATG ATG TAT ATT CAC CAC GGA TTA TCG GTG GGA	1707
Asp Gly Val Thr Ser Met Met Tyr Ile His His Gly Leu Ser Val Gly	
510 515 520	
TTC ACT GGG AAC TAC GAG GAA TAC TTT GGA CTC GCA ACT GAT GTG GAT	1755
Phe Thr Gly Asn Tyr Glu Glu Tyr Phe Gly Leu Ala Thr Asp Val Asp	
525 530 535	
GCT GTT GTG TAT CTG ATG CTG GTC AAC GAT CTT ATT CAT GGG CTT TTC	1803
Ala Val Val Tyr Leu Met Leu Val Asn Asp Leu Ile His Gly Leu Phe	
540 545 550	
CCA GAT GCA ATT ACC ATT GGT GAA GAT GTT AGC GGA ATG CCG ACA TTT	1851
Pro Asp Ala Ile Thr Ile Gly Glu Asp Val Ser Gly Met Pro Thr Phe	
555 560 565	
TGT ATT CCC GTC CAA GAG GGG GGT GTT GGC TTT GAC TAT CGG CTG CAT	1899
Cys Ile Pro Val Gln Glu Gly Gly Val Gly Phe Asp Tyr Arg Leu His	
570 575 580 585	

ATG	GCA	ATT	GCT	GAT	AAA	CGG	ATT	GAG	TTG	CTC	AAG	AAA	CGG	GAT	GAG	1947
Met	Ala	Ile	Ala	Asp	Lys	Arg	Ile	Glu	Leu	Leu	Lys	Lys	Arg	Asp	Glu	
				590				595						600		
GAT	TGG	AGA	GTG	GGT	GAT	ATT	GTT	CAT	ACA	CTG	ACA	AAT	AGA	AGA	TGG	1995
Asp	Trp	Arg	Val	Gly	Asp	Ile	Val	His	Thr	Leu	Thr	Asn	Arg	Arg	Trp	
			605				610						615			
TCG	GAA	AAG	TGT	GTT	TCA	TAC	GCT	GAA	AGT	CAT	GAT	CAA	GCT	CTA	GTC	2043
Ser	Glu	Lys	Cys	Val	Ser	Tyr	Ala	Glu	Ser	His	Asp	Gln	Ala	Leu	Val	
		620					625					630				
GGT	GAT	AAA	ACT	ATA	GCA	TTC	TGG	CTG	ATG	GAC	AAG	GAT	ATG	TAT	GAT	2091
Gly	Asp	Lys	Thr	Ile	Ala	Phe	Trp	Leu	Met	Asp	Lys	Asp	Met	Tyr	Asp	
	635					640					645					
TTT	ATG	GCT	CTG	GAT	AGA	CCG	TCA	ACA	TCA	TTA	ATA	GAT	CGT	GGG	ATA	2139
Phe	Met	Ala	Leu	Asp	Arg	Pro	Ser	Thr	Ser	Leu	Ile	Asp	Arg	Gly	Ile	
650					655					660				665		
GCA	TTG	CAC	AAG	ATG	ATT	AGG	CTT	GTA	ACT	ATG	GGA	TTA	GGA	GGA	GAA	2187
Ala	Leu	His	Lys	Met	Ile	Arg	Leu	Val	Thr	Met	Gly	Leu	Gly	Gly	Glu	
				670					675					680		
GGG	TAC	CTA	AAT	TTC	ATG	GGA	AAT	GAA	TTC	GGC	CAC	CCT	GAG	TGG	ATT	2235
Gly	Tyr	Leu	Asn	Phe	Met	Gly	Asn	Glu	Phe	Gly	His	Pro	Glu	Trp	Ile	
			685				690						695			
GAT	TTC	CCT	AGG	GCT	GAA	CAA	CAC	CTC	TCT	GAT	GGC	TCA	GTA	ATC	CCC	2283
Asp	Phe	Pro	Arg	Ala	Glu	Gln	His	Leu	Ser	Asp	Gly	Ser	Val	Ile	Pro	
		700				705						710				
GGA	AAC	CAA	TTC	AGT	TAT	GAT	AAA	TGC	AGA	CGG	AGA	TTT	GAC	CTG	GGA	2331
Gly	Asn	Gln	Phe	Ser	Tyr	Asp	Lys	Cys	Arg	Arg	Arg	Phe	Asp	Leu	Gly	
	715					720					725					
GAT	GCA	GAA	TAT	TTA	AGA	TAC	CGT	GGG	TTG	CAA	GAA	TTT	GAC	CGG	CCT	2379
Asp	Ala	Glu	Tyr	Leu	Arg	Tyr	Arg	Gly	Leu	Gln	Glu	Phe	Asp	Arg	Pro	
730					735				740					745		
ATG	CAG	TAT	CTT	GAA	GAT	AAA	TAT	GAG	TTT	ATG	ACT	TCA	GAA	CAC	CAG	2427
Met	Gln	Tyr	Leu	Glu	Asp	Lys	Tyr	Glu	Phe	Met	Thr	Ser	Glu	His	Gln	
				750					755					760		
TTC	ATA	TCA	CGA	AAG	GAT	GAA	GGA	GAT	AGG	ATG	ATT	GTA	TTT	GAA	AAA	2475
Phe	Ile	Ser	Arg	Lys	Asp	Glu	Gly	Asp	Arg	Met	Ile	Val	Phe	Glu	Lys	
			765				770						775			
GGA	AAC	CTA	GTT	TTT	GTC	TTT	AAT	TTT	CAC	TGG	ACA	AAA	AGC	TAT	TCA	2523
Gly	Asn	Leu	Val	Phe	Val	Phe	Asn	Phe	His	Trp	Thr	Lys	Ser	Tyr	Ser	
		780					785					790				
GAC	TAT	CGC	ATA	GCC	TGC	CTG	AAG	CCT	GGA	AAA	TAC	AAG	GTT	GCC	TTG	2571
Asp	Tyr	Arg	Ile	Ala	Cys	Leu	Lys	Pro	Gly	Lys	Tyr	Lys	Val	Ala	Leu	
	795					800					805					

46

GAC TCA GAT GAT CCA CTT TTT GGT GGC TTC GGG AGA ATT GAT CAT AAT 2619
 Asp Ser Asp Asp Pro Leu Phe Gly Gly Phe Gly Arg Ile Asp His Asn
 810 815 820 825

GCC GAA TAT TTC ACC TTT GAA GGA TGG TAT GAT GAT CGT CCT CGT TCA 2667
 Ala Glu Tyr Phe Thr Phe Glu Gly Trp Tyr Asp Asp Arg Pro Arg Ser
 830 835 840

ATT ATG GTG TAT GCA CCT TGT AAA ACA GCA GTG GTC TAT GCA CTA GTA 2715
 Ile Met Val Tyr Ala Pro Cys Lys Thr Ala Val Val Tyr Ala Leu Val
 845 850 855

GAC AAA GAA GAA GAA GAA GAA GAA GAA GAA GAA GAA GAA GAA GTA GCA GCA 2763
 Asp Lys Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Val Ala Ala
 860 865 870

GTA GAA GAA GTA GTA GTA GAA GAA GAA TGAACGAACT TGTGATCGCG 2810
 Val Glu Glu Val Val Val Glu Glu Glu
 875 880

TTGAAAGATT TGAACGCTAC ATAGAGCTTC TTGACGTATC TGGCAATATT GCATCAGTCT 2870

TGGCGGAATT TCATGTGACA CAAGGTTTGC AATTCTTTCC ACTATTAGTA GTGCAACGAT 2930

ATACGCAGAG ATGAAGTGCT GAACAAACAT ATGTAAAATC GATGAATTTA TGTCGAATGC 2990

TGGGACGATC GAATTCCTGC AGGCCGGGGG ACCCCTTAGT TCT 3033

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 882 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

Met Val Tyr Thr Leu Ser Gly Val Arg Phe Pro Thr Val Pro Ser Val
 1 5 10 15

Tyr Lys Ser Asn Gly Phe Ser Ser Asn Gly Asp Arg Arg Asn Ala Asn
 20 25 30

Val Ser Val Phe Leu Lys Lys His Ser Leu Ser Arg Lys Ile Leu Ala
 35 40 45

Glu Lys Ser Ser Tyr Asn Ser Glu Phe Arg Pro Ser Thr Val Ala Ala
 50 55 60

Ser Gly Lys Val Leu Val Pro Gly Thr Gln Ser Asp Ser Ser Ser Ser
 65 70 75 80

Ser Thr Asp Gln Phe Glu Phe Thr Glu Thr Ser Pro Glu Asn Ser Pro
 85 90 95

Ala Ser Thr Asp Val Asp Ser Ser Thr Met Glu His Ala Ser Gln Ile
 100 105 110
 Lys Thr Glu Asn Asp Asp Val Glu Pro Ser Ser Asp Leu Thr Gly Ser
 115 120 125
 Val Glu Glu Leu Asp Phe Ala Ser Ser Leu Gln Leu Gln Glu Gly Gly
 130 135 140
 Lys Leu Glu Glu Ser Lys Thr Leu Asn Thr Ser Glu Glu Thr Ile Ile
 145 150 155 160
 Asp Glu Ser Asp Arg Ile Arg Glu Arg Gly Ile Pro Pro Pro Gly Leu
 165 170 175
 Gly Gln Lys Ile Tyr Glu Ile Asp Pro Leu Leu Thr Asn Tyr Arg Gln
 180 185 190
 His Leu Asp Tyr Arg Tyr Ser Gln Tyr Lys Lys Leu Arg Glu Ala Ile
 195 200 205
 Asp Lys Tyr Glu Gly Gly Leu Glu Ala Phe Ser Arg Gly Tyr Glu Lys
 210 215 220
 Met Gly Phe Thr Arg Ser Ala Thr Gly Ile Thr Tyr Arg Glu Trp Ala
 225 230 235 240
 Leu Gly Ala Gln Ser Ala Ala Leu Ile Gly Asp Phe Asn Asn Trp Asp
 245 250 255
 Ala Asn Ala Asp Ile Met Thr Arg Asn Glu Phe Gly Val Trp Glu Ile
 260 265 270
 Phe Leu Pro Asn Asn Val Asp Gly Ser Pro Ala Ile Pro His Gly Ser
 275 280 285
 Arg Val Lys Ile Arg Met Asp Thr Pro Ser Gly Val Lys Asp Ser Ile
 290 295 300
 Pro Ala Trp Ile Asn Tyr Ser Leu Gln Leu Pro Asp Glu Ile Pro Tyr
 305 310 315 320
 Asn Gly Ile His Tyr Asp Pro Pro Glu Glu Glu Arg Tyr Ile Phe Gln
 325 330 335
 His Pro Arg Pro Lys Lys Pro Lys Ser Leu Arg Ile Tyr Glu Ser His
 340 345 350
 Ile Gly Met Ser Ser Pro Glu Pro Lys Ile Asn Ser Tyr Val Asn Phe
 355 360 365
 Arg Asp Glu Val Leu Pro Arg Ile Lys Lys Leu Gly Tyr Asn Ala Leu
 370 375 380
 Gln Ile Met Ala Ile Gln Glu His Ser Tyr Tyr Ala Ser Phe Gly Tyr
 385 390 395 400

His	Val	Thr	Asn	Phe	Phe	Ala	Pro	Ser	Ser	Arg	Phe	Gly	Thr	Pro	Asp
				405					410					415	
Asp	Leu	Lys	Ser	Leu	Ile	Asp	Lys	Ala	His	Glu	Leu	Gly	Ile	Val	Val
			420					425					430		
Leu	Met	Asp	Ile	Val	His	Ser	His	Ala	Ser	Asn	Asn	Thr	Leu	Asp	Gly
		435					440					445			
Leu	Asn	Met	Phe	Asp	Cys	Thr	Asp	Ser	Cys	Tyr	Phe	His	Ser	Gly	Ala
	450					455					460				
Arg	Gly	Tyr	His	Trp	Met	Trp	Asp	Ser	Arg	Leu	Phe	Asn	Tyr	Gly	Asn
465					470					475					480
Trp	Glu	Val	Leu	Arg	Tyr	Leu	Leu	Ser	Asn	Ala	Arg	Trp	Trp	Leu	Asp
				485					490					495	
Ala	Phe	Lys	Phe	Asp	Gly	Phe	Arg	Phe	Asp	Gly	Val	Thr	Ser	Met	Met
			500					505					510		
Tyr	Ile	His	His	Gly	Leu	Ser	Val	Gly	Phe	Thr	Gly	Asn	Tyr	Glu	Glu
		515					520					525			
Tyr	Phe	Gly	Leu	Ala	Thr	Asp	Val	Asp	Ala	Val	Val	Tyr	Leu	Met	Leu
	530					535					540				
Val	Asn	Asp	Leu	Ile	His	Gly	Leu	Phe	Pro	Asp	Ala	Ile	Thr	Ile	Gly
545					550					555					560
Glu	Asp	Val	Ser	Gly	Met	Pro	Thr	Phe	Cys	Ile	Pro	Val	Gln	Glu	Gly
				565					570					575	
Gly	Val	Gly	Phe	Asp	Tyr	Arg	Leu	His	Met	Ala	Ile	Ala	Asp	Lys	Arg
			580					585					590		
Ile	Glu	Leu	Leu	Lys	Lys	Arg	Asp	Glu	Asp	Trp	Arg	Val	Gly	Asp	Ile
		595					600					605			
Val	His	Thr	Leu	Thr	Asn	Arg	Arg	Trp	Ser	Glu	Lys	Cys	Val	Ser	Tyr
	610					615					620				
Ala	Glu	Ser	His	Asp	Gln	Ala	Leu	Val	Gly	Asp	Lys	Thr	Ile	Ala	Phe
625					630					635					640
Trp	Leu	Met	Asp	Lys	Asp	Met	Tyr	Asp	Phe	Met	Ala	Leu	Asp	Arg	Pro
				645					650					655	
Ser	Thr	Ser	Leu	Ile	Asp	Arg	Gly	Ile	Ala	Leu	His	Lys	Met	Ile	Arg
			660					665					670		
Leu	Val	Thr	Met	Gly	Leu	Gly	Gly	Glu	Gly	Tyr	Leu	Asn	Phe	Met	Gly
		675					680					685			
Asn	Glu	Phe	Gly	His	Pro	Glu	Trp	Ile	Asp	Phe	Pro	Arg	Ala	Glu	Gln
	690					695					700				

His Leu Ser Asp Gly Ser Val Ile Pro Gly Asn Gln Phe Ser Tyr Asp
 705 710 715 720
 Lys Cys Arg Arg Arg Phe Asp Leu Gly Asp Ala Glu Tyr Leu Arg Tyr
 725 730 735
 Arg Gly Leu Gln Glu Phe Asp Arg Pro Met Gln Tyr Leu Glu Asp Lys
 740 745 750
 Tyr Glu Phe Met Thr Ser Glu His Gln Phe Ile Ser Arg Lys Asp Glu
 755 760 765
 Gly Asp Arg Met Ile Val Phe Glu Lys Gly Asn Leu Val Phe Val Phe
 770 775 780
 Asn Phe His Trp Thr Lys Ser Tyr Ser Asp Tyr Arg Ile Ala Cys Leu
 785 790 795 800
 Lys Pro Gly Lys Tyr Lys Val Ala Leu Asp Ser Asp Asp Pro Leu Phe
 805 810 815
 Gly Gly Phe Gly Arg Ile Asp His Asn Ala Glu Tyr Phe Thr Phe Glu
 820 825 830
 Gly Trp Tyr Asp Asp Arg Pro Arg Ser Ile Met Val Tyr Ala Pro Cys
 835 840 845
 Lys Thr Ala Val Val Tyr Ala Leu Val Asp Lys Glu Glu Glu Glu Glu
 850 855 860
 Glu Glu Glu Glu Glu Glu Val Ala Ala Val Glu Glu Val Val Val Glu
 865 870 875 880
 Glu Glu

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2576 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TCATTAAAGA GGAGAAATTA ACTATGAGAG GATCTCACCA TCACCATCAC CATGGGATCT	60
TGGCTGAAAA GTCTTCTTAC AATTCCGAAT TCCGACCTTC TACAGTTGCA GCATCGGGGA	120
AAGTCCTTGT GCCTGGAACC CAGAGTGATA GCTCCTCATC CTCAACAAAC CAATTTGAGT	180
TCACTGAGAC ATCTCCAGAA AATTCCCCAG CATCAACTGA TG TAGATAGT TCAACAATGG	240
AACACGCTAG CCAGATTAAA ACTGAGAACG ATGACGTTGA GCCGTCAAGT GATCTTACAG	300

GAAGTGTGA AGAGCTGGAT TTTGCTTCAT CACTACAAC	360
ACAAGAAGGT GGTAACTGG	
AGGAGTCTAA AACATTAAAT ACTTCTGAAG AGACAATTAT	420
TGATGAATCT GATAGGATCA	
GAGAGAGGGG CATCCCTCCA CCTGGACTTG GTCAGAAGAT	480
TTATGAAATA GACCCCTTT	
TGACAACTA TCGTCAACAC CTTGATTACA GGTATTCACA	540
GTACAAGAAA CTGAGGGAGG	
CAATTGACAA GTATGAGGGT GGTTTGGAAG CTTTTCTCG	600
TGGTTATGAA AAAATGGGTT	
TCACTCGTAG TGCTACAGGT ATCACTTACC GTGAGTGGGC	660
TCCTGGTGCC CAGTCAGCTG	
CCCTCATTGG AGATTTCAAC AATTGGGACG CAAATGCTGA	720
CATTATGACT CGGAATGAAT	
TTGGTGTCTG GGAGATTTTT CTGCCAAATA ATGTGGATGG	780
TTCTCCTGCA ATTCCTCATG	
GGTCCAGAGT GAAGATACGT ATGGACACTC CATCAGGTGT	840
TAAGGATTCC ATTCCTGCTT	
GGATCACTA CTCTACAGCT TCCTGATGAA ATTCCATATA	900
ATGGAATATA TTATGATCCA	
CCCGAAGAGG AGAGGTATAT CTTCCAACAC CCACGGCCAA	960
AGAAACCAA GTCGCTGAGA	
ATATATGAAT CTCATATTGG AATGAGTAGT CCGGAGCCTA	1020
AAATTAATC ATACGTGAAT	
TTTAGAGATG AAGTTCTTCC TCGCATAAAA AAGCTTGGGT	1080
ACAATGCGCT GCAAATTATG	
GCTATTCAAG AGCATTCTTA TTATGCTAGT TTTGGTTATC	1140
ATGTCACAAA TTTTTTGCA	
CCAAGCAGCC GTTTTGGAAC GCCCACGAC CTTAAGTCTT	1200
TGATTGATAA AGCTCATGAG	
CTAGGAATTG TTGTTCTCAT GGACATTGTT CACAGCCATG	1260
CATCAAATAA TACTTTAGAT	
GGACTGAACA TGTTTGACGG CACCGATAGT TGTTACTTTC	1320
ACTCTGGAGC TCGTGGTTAT	
CATTGGATGT GGGATTCCCG CCTTTTAAAC TATGGAACT	1380
GGGAGGTACT TAGGTATCTT	
CTCTCAAATG CGAGATGGTG GTTGGATGAG TTCAAATTTG	1440
ATGGATTTAG ATTTGATGGT	
GTGACATCAA TGATGTATAC TCACCACGGA TTATCGGTGG	1500
GATTCACTGG GAACTACGAG	
GAATACTTTG GACTCGCAAC TGATGTGGAT GCTGTTGTGT	1560
ATCTGATGCT GGTCAACGAT	
CTTATTCATG GGCTTTTCCC AGATGCAATT ACCATTGGTG	1620
AAGATGTTAG CGGAATGCCG	
ACATTTTGTA TTCCCGTTCA AGATGGGGGT GTTGGCTTTG	1680
ACTATCGGCT GCATATGGCA	
ATTGCTGATA AATGGATTGA GTTGCTCAAG AAACGGGATG	1740
AGGATTGGAG AGTGGGTGAT	
ATTGTTCATA CACTGACAAA TAGAAGATGG TCGGAAAAGT	1800
GTGTTTCATA CGCTGAAAGT	
CATGATCAAG CTCTAGTCGG TGATAAACT ATAGCATTCT	1860
GGCTGATGGA CAAGGATATG	
TATGATTTTA TGGCTCTGGA TAGACCGCCA ACATCATTAA	1920
TAGATCGTGG GATAGCATTG	
CACAAGATGA TTAGGCTTGT AACTATGGGA TTAGGAGGAG	1980
AAGGGTACCT AAATTTTCATG	

GGAAATGAAT TCGGCCACCC TGAGTGGATT GATTTCCCTA GGGCTGAACA ACACCTCTCT	2040
GATGACTCAG TAATTCCCGG AAACCAATTC AGTTATGATA AATGCAGACG GAGATTTGAC	2100
CTGGGAGATG CAGAATATTT AAGATACCGT GGGTTGCAAG AATTTGACCG GGCTATGCAG	2160
TATCTTGAAG ATAAATATGA GTTTATGACT TCAGAACACC AGTTCATATC ACGAAAGGAT	2220
GAAGGAGATA GGATGATTGT ATTTGAAAAA GGAAACCTAG TTTTGTCTT TAATTTTCAC	2280
TGGACAAAAA GCTATTCAGA CTATCGCATA GGCTGCCTGA AGCCTGGAAA ATACAAGGTT	2340
GCCTTGACT CAGATGATCC ACTTTTTGGT GGCTTCGGGA GAATTGATCA TAATGCCGAA	2400
TATTTACCT TTGAAGGATG GTATGATGAT CGTCCTCGTT CAATTATGGT GTATGCACCT	2460
TGTAGAACAG CAGTGGTCTA TGCACTAGTA GACAAAGAAG AAGAAGAAGA AGAAGAAGAA	2520
GAAGAAGTAG CAGTAGTAGA AGAAGTAGTA GTAGAAGAAG AATGAACGAA CTTGTG	2576

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2529 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

GGATGCTAAT GTTCTGTAT TCTTGAAAAA GCACTCTCTT TCACGGAAGA TCTTGGCTGA	60
AAAGTCTTCT TACAATTCCG AATCCCGACC TTCTACAGTT GCAGCATCGG GGAAAGTCCT	120
TGTGCCTGGA AYCCAGAGTG ATAGCTCCTC ATCCTCAACA GACCAATTTG AGTTCACTGA	180
GACATCTCCA GAAAATTCCC CAGCATCAAC TGATGTAGAT AGTTCAACAA TGGAACACGC	240
TAGCCAGATT AAAACTGAGA ACGATGACGT TGAGCCGTCA AGTGATCTTA CAGGAAGTGT	300
TGAAGAGCTG GATTTTGCTT CATCACTACA ACTACAAGAA GGTGGTAAAC TGGAGGAGTC	360
TAAACATTA AATACTTCTG AAGAGACAAT TATTGATGAA TCTGATAGGA TCAGAGAGAG	420
GGGCATCCCT CCACCTGGAC TTGGTCAGAA GATTTATGAA ATAGACCCCC TTTTGACAAA	480
CTATCGTCAA CACCTTGATT ACAGGTATTC ACAGTACAAG AAACTGAGGG AGGCAATTGA	540
CAAGTATGAG GGTGGTTTGG AAGCTTTTTT TCGTGGTTAT GAAAAAATGG GTTTCACCTG	600
TAGTGCTACA GGTATCACTT ACCGTGAGTG GGCTCCTGGT GCCCAGTCAG CTGCCCTCAT	660
TGGAGATTTT AACAATTGGG ACGCAAATGC TGACATTATG ACTCGGAATG AATTTGGTGT	720
CTGGGAGATT TTTCTGCCAA ATAATGTGGA TGGTTCTCCT GCAATTCCTC ATGGGTCCAG	780

AGTGAAGATA CGYATGGACA CTCCATCAGG TGTTAAGGAT TCCATTCCTG CTTGGATCAA 840
CTACTCTTTA CAGCTTCCTG ATGAAATTCC ATATAATGGA ATATATTATG ATCCACCCGA 900
AGAGGAGAGG TATRTCTTCC AACACCCACG GCCAAAGAAA CCAAAGTCGC TGAGAATATA 960
TGAATCTCAT ATTGGAATGA GTAGTCCGGA GCCTAAAATT AACTCATACG TGAATTTTAG 1020
AGATGAAGTT CTTCTCGCA TAAAAAASCT TGGGTACAAT GCGGTGCAAA TTATGGCTAT 1080
TCAAGAGCAT TCTTATTATG CTAGTTTTGG TTATCATGTC ACAAATTTTT TTGCACCAAG 1140
CAGCCGTTTT GGAACGCCCG ACGACCTTAA GTCTTTGATT GATAAAGCTC ATGAGCTAGG 1200
AATTGTTGTT CTCATGGACA TTGTTACAG CCATGCATCA AATAATACTT TAGATGGACT 1260
GAACATGTTT GACGGCACAG ATAGTTGTTA CTTTCACTCT GGAGCTCGTG GTTATCATTG 1320
GATGTGGGAT TCCCGCCTCT TTAACATGGA AAAGTGGGAG GTACTTAGGT ATCTTCTCTC 1380
AAATGCGAGA TGGTGGTTGG ATGAGTTCAA ATTTGATGGA TTTAGATTTG ATGGTGTGAC 1440
ATCAATGATG TATACTCACC ACGGATTATC GGTGGGATTC ACTGGGAACT ACGAGGAATA 1500
CTTTGGACTC GCAACTGATG TGGATGCTGT TGTGTATCTG ATGCTGGTCA ACGATCTTAT 1560
TCACGGGCTT TTCCCAGATG CAATTACCAT TGGTGAAGAT GTTAGCGGAA TGCCGACATT 1620
TTGTATTCCC GTTCAAGATG GGGGTGTTGG CTTTGACTAT CGGCTGCATA TGGCAATTGC 1680
TGATAAATGG ATTGAGTTGC TCAAGAAACG GGATGAGGAT TGGAGAGTGG GTGATATTGT 1740
TCATACACTG ACAAATAGAA GATGGTCGGA AAAGTGTGTT TCATMCGCTG AAAGTCATGA 1800
TCAAGCTCTA GTCGGTGATA AAAGTATAGC ATYCTGGCTG ATGGACAAGG ATATGTATGA 1860
TTTTATGGCT CTGGATAGAC CGYCAACAYC ATTAATAGAT CGTGGGATAG CATTGCACAA 1920
GATGATTAGG CTTGTAAC TAAGGATTAGG AGGAGAAGGG TACCTAAATT TCATGGGAAA 1980
TGAATTCGGC CACCCTGAGT GGATTGATTT CCCTAGGGCT GARCAACACC TCTCTGATGG 2040
CTCAGTAATT CCCGGAAACC AATTCAGTTA TGATAAATGC AGACGGAGAT TTGACCTGGG 2100
AGATGCAGAA TATTTAAGAT ACCATGGGTT GCAAGAATTT GACCGGGCTA TGCAGTATCT 2160
TGAAGATAAA TATGAGTTTA TGAATTCAGA ACACCAAGTT ATATCACGAA AGGATGAAGG 2220
AGATAGGATG ATTGTATTTG AAARAGGAAA CCTAGTTTTT GTCTTTAATT TTTACTGGAC 2280
AAATAGCTAT TCAGACTATC GCATAGGCTG CCTGAAGCCT GGAAAATACA AGGTTGGCTT 2340
GGAATCAGAT GATCCACTTT TTGGTGGCTT CGGGAGAATT GATCATAATG CCGAATATTT 2400
CACCTCTGAA GGATCGTATG ATGATCGTCC TCGTTCAATT ATGGTGTATG CACCTAGTAG 2460

AACAGCAGTG GTCTATGCAC TAGTAGACAA ANTAGAAGNA GAAGAAGAAG TAGAANCCGN 2520
 NGAAGAATT 2529

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3231 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GATTTAATAC GACTCACTAT AGGGATTTTT TTTTTTTTTT TTTTAAAAC CTCCTCCACT 60
 CAGTCTTGGG ATCTCTCTCT CTCTTCACGC TTCTCTTGGG GCCTTGAAC CAGCAATTTG 120
 AACTCAGTT AGTTAACTC CTATCACTCA TCAGATCTCT ATTTTTTCTC TTAATTCCAA 180
 CCAAGGAATG AATTTAAAGA TTAGATTGA AGGAGAGAAG AAGAAAGATG GTGTATACAC 240
 TCTCTGGAGT TCGTTTTCT ACTGTTCCAT CAGTGTACAA ATCTAATGGA TTCAGCAGTA 300
 ATGGTGATCG GAGGAATGCT AATGTTTCTG TATTCTTGAA AAAGCACTCT CTTTCACGGA 360
 AGATCTTGGC TGAAAAGTCT TCTTACGATT CCGAATCCCG ACCTTCTACA GTTGACGAT 420
 CGGGGAAAGT CTTGTACCT GGAATCCAGA GTGATAGCTC CTCATCCTCA ACAGACCAAT 480
 TTGAGTTCAC TGAGACAGCT CCAGAAAATT CCCCAGCATC AACTGATGTG GATAGTTCAA 540
 CAATGGAACA CGCTAGCCAG ATTTAAACTG AGAACGATGA CGTTGAGCCG TCAAGTGATC 600
 TTACAGGAAG TGTTGAAGAG TTGGATTTTG CTTCATCACT ACAACTACAA GAAGGTGGTA 660
 AACTGGAGGA GTCTAAAACA TTAAATACTT CTGAAGAGAC AATTATTGAT GAATCTGATA 720
 GGATCAGAGA GAGGGGCATC CCTCCACCTG GACTTGGTCA GAAGATTTAT GAAATAGACC 780
 CCCTTTTGAC AACTATCGT CAACACCTTG ATTACAGGTA TTCACAGTAC AAGAAAATGA 840
 GGGAGGCAAT TGACAAGTAT GAGGGTGGTT TGGAAGCTTT TTCTCGTGGT TATGAAAAAA 900
 TGGGTTTCAC TCGTAGTGCT ACAGGTATCA CTTACCGTGA GTGGGCTCCT GGTGCCCAGT 960
 CAGCTGCTCT CATTGGAGAT TTCAACAATT GGGACGCAAA TGCTGACATT ATGACTCGGA 1020
 ATGAATTTGG TGTCTGGGAG ATTTTTCTGC CAAATAATGT GGATGGTTCT CTGCAATTC 1080
 CTCATGGGTC CAGAGTGAAG ATACGCATGG AACTTCATC AGGTGTTAAG GATTCCATTC 1140
 CTGCTTGGAT CAACTACTCT TTACAGCTTC CTGATGAAAT TCCATATAAT GGAATATATT 1200
 ATGATCCACC CGAAGAGGAG AGGTATGTCT TCCAACACCC ACGGCCAAAG AAACCAAAGT 1260

CGCTGAGAAT ATATGAATCT CATATTGGAA TGAGTAGTCC GGAGCCTAAA ATTAACATCAT 1320
ACGTGAATTT TAGAGATGAA GTTCTTCCTC GCATAAAAAA CCTTGGGTAC AATGCGGTGC 1380
AAATTATGGC TATTCAAGAG CATTCTTATT ATGCTAGTTT TGGTTATCAT GTCACAAATT 1440
TTTTTGCACC AAGCAGCCGT TTTGGAACGC CCGACGACCT TAAGTCTTTG ATTGATAAAG 1500
CTCATGAGCT AGGAATTGTT GTTCTCATGG ACATTGTTCA CAGCCATGCA TCAAATAATA 1560
CTTTAGATGG ACTGAACATG TTTGACGGCA CAGATAGTTG TTACTTTTAC TCTGGAGCTC 1620
GTGGTTATCA TTGGATGTGG GATTCCCGCC TCTTTAACTA TGGAAACTGG GAGGTACTTA 1680
GGTATCTTCT CTCAAATGCG AGATGGTGGT TGGATGAGTG CAAATTTGRT GGATTTAGAT 1740
TTGATGGTGT GACATCAATG ATGTATACTC ACCACGGATT ATCGGTGGGA TTCACTGGGA 1800
ACTACGAGGA ATACTTTGGA CTCGCAACTG ATGTRGATGC TGCCGTGTAT CTGATGCTGG 1860
CCAACGATCT TATTCATGGG CTTTTCCAG ATGCAATTAC CATTGGTGAA GATGTTAGCG 1920
GAATGCCGAC ATTTTGTATT CCCGTTCAAG ATGGGGGTGT TGGCTTTGAC TATCGGCTGC 1980
ATATGGCAAT TGCTGATAAA TGGATTGAGT TGCTCAAGAA ACGGGATGAG GATTGGAGAG 2040
TGGGTGATAT TGTTCATACA CTGACAAATA GAAGATGGTC GGAAAAGTGT GTTTCATACG 2100
CTGAAAGTCA TGATCAAGCT CTAGTCGGTG ATAAACTAT AGCATTCTGG CTGATGGACA 2160
AGGATATGTA TGATTTTATG GCTTTGGATA GACCGTCAAC ATCATTAAATA GATCGTGGGA 2220
TAGCATTGCA CAAGATGATT AGGCTTGTA CTATGGGATT AGGAGGAGAA GGGTACCTAA 2280
ATTTTCATGGG AAATGAATTC GGCCACCCTG AGTGGATTGA TTTCCCTAGG GCTGAACAAC 2340
ACCTCTCTGA TGGCTCAGTA ATTCCCGGAA ACCAATTCAG TTATGATAAA TGCAGACGGA 2400
GATTTGACCT GGGAGATGCA GAATATTTAA GATACCGTGG GTTGCAAGAA TTTGACCGGG 2460
CTATGCAGTA TCTTGAAGAT AAATATGAGT TTATGACTTC AGAACACCAG TTCATATCAC 2520
GAAAGGATGA AGGAGATAGG ATGATTGTAT TTGAAAAAGG AAACCTAGTT TTTGTCTTTA 2580
ATTTTCACTG GACAAAAAGC TATTCAGACT ATCGCATAGG CTGGCTGAAG CCTGGAAAAT 2640
ACAAGGTTGC CTTGGACTCA GATGATCCAC TTTTGGTGG CTTCGGGAGA ATTGATCATA 2700
ATGCCGAATG TTTACCTTT GAAGGATGGT ATGATGATCG TCCTCGTTCA ATTATGGTGT 2760
ATGCACCTAG TAGAACAGCA GTGGTCTATG CACTAGTAGA CAAAGAAGAA GAAGAAGAAG 2820
AAGTAGCAGT AGTAGAAGAA GTAGTAGTAG AAGAAGAATG AACGAACCTG TGATCGCGTT 2880
GAAAGATTTG AACGCTACAT AGAGCTTCTT GACGTATCTG GCAATATTGC ATCAGTCTTG 2940

GCGGAATTC ATGTGACAAA AGGTTTGCAA TTCTTCCAC TATTAGTAGT GCAACGATAT	3000
ACGCAGAGAT GAAGTGCTGA ACAAACATAT GTAAAATCGA TGAATTTATG TCGAATGCTG	3060
GGACGGGCTT CAGCAGGTTT TGCTTAGTGA GTTCTGTAAA TTGTCATCTC TTTANATGTA	3120
GAGCCCACTA GAAATCAATT ATGTGAGACC TAAAAACAA TAACCATAAA ATGGAAATAG	3180
TGCTGATCTA ATGATGTTTT AANCCNNNA AAAAAAAAAA AAAAATCTGA G	3231

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2578 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

TCATTAAAGA GGAGAAATTA ACTATGAGAG GATCTCACCA TCACCATCAC CATGGGATCT	60
TGGCTGAAAA GTCTTCTTAC AATTCCGAAT TCCGACCTTC TACAGTTGCA GCATCGGGGA	120
AAGTCCTTGT GCCTGGAACC CAGAGTGATA GCTCCTCATC CTCAACAAAC CAATTTGAGT	180
TCACTGAGAC ATCTCCAGAA AATCCCCAG CATCAACTGA TGATAGATAGT TCAACAATGG	240
AACACGCTAG CCAGATTAAG ACTGAGAACG ATGACGTTGA GCCGTCAAGT GATCTTACAG	300
GAAGTGTTGA AGAGCTGGAT TTTGCTTCAT CACTACAAC ACAAGAAGGT GGTAAACTGG	360
AGGAGTCTAA AACATTAAAT ACTTCTGAAG AGACAATTAT TGATGAATCT GATAGGATCA	420
GAGAGAGGGG CATCCCTCCA CCTGGACTTG GTCAGAAGAT TTATGAAATA GACCCCTTT	480
TGACAAACTA TCGTCAACAC CTTGATTACA GGTATTCACA GTACAAGAAA CTGAGGGAGG	540
CAATTGACAA GTATGAGGGT GGTTTGGAAG CTTTTTCTCG TGGTTATGAA AAAATGGGTT	600
TCACTCGTAG TGCTACAGGT ATCACTTACC GTGAGTGGGC TCCTGGTGCC CAGTCAGCTG	660
CCCTCATTGG AGATTCAAC AATTGGGACG CAAATGCTGA CATTATGACT CGGAATGAAT	720
TTGGTGTCTG GGAGATTTTT CTGCCAAATA ATGTGGATGG TTCTCCTGCA ATTCCTCATG	780
GGTCCAGAGT GAAGATACGT ATGGACACTC CATCAGGTGT TAAGGATTCC ATTCCTGCTT	840
GGATCAACTA CTCTCACAG CTTCTGATG AAATTCCATA TAATGGAATA TATTATGATC	900
CACCCGAAGA GGAGAGGTAT ATCTTCCAAC ACCCACGGCC AAAGAAACCA AAGTCGCTGA	960
GAATATATGA ATCTCATATT GGAATGAGTA GTCCGGAGCC TAAATTAAC TCATACGTGA	1020
ATTTTAGAGA TGAAGTTCTT CCTCGCATAA AAAAGCTTGG GTACAATGCG GTGCAAATTA	1080

TGGCTATTCA AGAGCATTCT TATTATGCTA GTTTTGGTTA TCATGTCACA AATTTTTTTG	1140
CACCAAGCAG CCGTTTTGGA ACGCCCGACG ACCTTAAGTC TTTGATTGAT AAAGCTCATG	1200
AGCTAGGAAT TGTTGTTCTC ATGGACATTG TTCACAGCCA TGCATCAAAT AATACTTTAG	1260
ATGGACTGAA CATGTTTGAC GGCACCGATA GTTGTTACTT TCACTCTGGA GCTCGTGGTT	1320
ATCATTGGAT GTGGGATTCC CGCCTTTTTA ACTATGGAAA CTGGGAGGTA CTTAGGTATC	1380
TTCTCTCAA TGCAGATGG TGGTTGGATG AGTTCAAATT TGATGGATTT AGATTTGATG	1440
GTGTGACATC AATGATGTAT ACTCACCACG GATTATCGGT GGGATTCACT GGGAACTACG	1500
AGGAATACTT TGGACTCGCA ACTGATGTGG ATGCTGTTGT GTATCTGATG CTGGTCAACG	1560
ATCTTATTCA TGGGCTTTTC CCAGATGCAA TTACCATTGG TGAAGATGTT AGCGGAATGC	1620
CGACATTTTG TATTCCCGTT CAAGATGGGG GTGTTGGCTT TGAATATCGG CTGCATATGG	1680
CAATTGCTGA TAAATGGATT GAGTTGCTCA AGAAACGGGA TGAGGATTGG AGAGTGGGTG	1740
ATATTGTTCA TACTGACA AATAGAAGAT GGTCGGAAAA GTGTGTTTCA TACGCTGAAA	1800
GTCATGATCA AGCTCTAGTC GGTGATAAAA CTATAGCATT CTGGCTGATG GACAAGGATA	1860
TGTATGATTT TATGGCTCTG GATAGACCGC CAACATCATT AATAGATCGT GGGATAGCAT	1920
TGCACAAGAT GATTAGGCTT GTAACATATG GATTAGGAGG AGAAGGGTAC CTAAATTTCA	1980
TGGGAAATGA ATTCGGCCAC CCTGAGTGGA TTGATTTCCC TAGGGCTGAA CAACACCTCT	2040
CTGATGACTC AGTAATTTCC GGAAACCAAT TCAGTTATGA TAAATGCAGA CGGAGATTTG	2100
ACCTGGGAGA TGCAGAATAT TTAAGATACC GTGGGTTGCA AGAATTTGAC CGGGCTATGC	2160
AGTATCTTGA AGATAAATAT GAGTTTATGA CTTCAGAACA CCAGTTCATA TCACGAAAGG	2220
ATGAAGGAGA TAGGATGATT GTATTTGAAA AAGGAAACCT AGTTTTTGTC TTTAATTTTC	2280
ACTGGACAAA AAGCTATTCA GACTATCGCA TAGGCTGCCT GAAGCCTGGA AAATACAAGG	2340
TTGCCTTGGA CTCAGATGAT CCACTTTTTG GTGGCTTCGG GAGAATTGAT CATAATGCCG	2400
AATATTTTCA CTTTGAAGGA TGGTATGATG ATCGTCCTCG TTCAATTATG GTGTATGCAC	2460
CTTGTAAGAC AGCAGTGGTC TATGCACTAG TAGACAAAGA AGAAGAAGAA GAAGAAGAAG	2520
AAGAAGAAGT AGCAGTAGTA GAAGAAGTAG TAGTAGAAGA AGAATGAACG AACTTGTG	2578

57

(2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AATTTYATGG GNAAYGARTT YGG

23

CLAIMS

1. Starch extracted from a potato plant and having an amylose content of at least 35%, as judged by the iodometric assay method of Morrison & Laignelet (1983 J. Cereal Science 1, 9-20).
2. Starch according to claim 1, having an amylose content of at least 37%, as judged by the method defined in claim 1.
3. Starch according to claim 1, having an amylose content of at least 40%, as judged by the method defined in claim 1.
4. Starch according to claim 1, having an amylose content of at least 50%, as judged by the method defined in claim 1.
5. Starch according to claim 1, having an amylose content of at least 66%, as judged by the method defined in claim 1.
6. Starch according to any one of claims 1-5, having an amylose content of 35 - 66%, as judged by the method defined in claim 1.
7. Starch which as extracted from a potato plant by wet milling at ambient temperature has a viscosity onset temperature in the range 70 - 95°C, as judged by viscoamylograph of a 10% w/w aqueous suspension thereof, performed at atmospheric pressure using the Newport Scientific Rapid Visco Analyser 3C with a heating profile of holding at 50°C for 2 minutes, heating from 50 to 95°C at a rate of 1.5°C per minute, holding at 95°C for 15 minutes, cooling from 95 to 50°C at a rate of 1.5°C per minute, and then holding at 50°C for 15 minutes.
8. Starch which as extracted from a potato plant by wet milling at ambient temperature has peak viscosity in the range 500 - 12 stirring number units (SNU), as judged by viscoamylograph conducted according to the protocol defined in claim 7.

9. Starch which as extracted from a potato plant by wet milling at ambient temperature has a pasting viscosity in the range 214 - 434 SNUs, as judged by viscoamylograph conducted according to the protocol defined in claim 7.
10. Starch which as extracted from a potato plant by wet milling at ambient temperature has a set-back viscosity in the range 450 - 618 SNUs, as judged by viscoamylograph conducted according to the protocol defined in claim 7.
11. Starch which as extracted from a potato plant by wet milling at ambient temperature has a set-back viscosity in the range 14 - 192 SNUs, as judged by viscoamylograph conducted according to the protocol defined in claim 7.
12. Starch which as extracted from a potato plant by wet milling at ambient temperature has a peak viscosity in the range 200 - 500 SNUs and a set-back viscosity in the range 275-618 SNUs as judged by viscoamylograph according to the protocol defined in claim 7.
13. Starch which as extracted from a potato plant by wet milling at ambient temperature has a viscosity which does not decrease between the start of the heating phase (step 2) and the start of the final holding phase (step 5) and has a set-back viscosity of 303 SNUs or less as judged by viscoamylograph according to the protocol defined in claim 7.
14. Starch which as extracted from a potato plant by wet milling at ambient temperature displays no significant increase in viscosity as judged by viscoamylograph conducted according to the protocol defined in claim 7.
15. Starch which as extracted from a potato plant by wet milling at ambient temperature, is in accordance with claim 7 and in accordance with any one of claims 8 to 14.
16. Starch according to any one of claims 7 to 15, having an amylose content in the range 35 - 66%, as judged by the method of Morrison & Laignelet defined in claim 1.

17. Starch which is extracted from a potato plant, has a phosphorus content in excess of 200mg/100grams dry weight starch.
18. Starch according to claim 17, having a phosphorus content in the range 200 - 240mg/100grams dry weight starch.
19. Starch according to claim 17 or 18, further in accordance with any one of claims 1 to 16.
20. Starch prepared by physical, chemical and/or enzymatic treatment of a starch initially having properties in accordance with any one of claims 1-19.
21. Starch according to claim 20, being resistant starch prepared by physical, chemical and/or enzymatic treatment of a starch initially having properties in accordance with any one of claims 1-19.
22. Starch according to claim 21, comprising in excess of 5% total dietary fibre, as determined according to the method of Prosky *et al.*, (1985 J. Assoc. Off. Anal. Chem. 68, 677).
23. Use of starch according to any one of claims 1-22 in the preparation or processing of a foodstuff.
24. Use of starch according to claim 23, wherein the starch is used to provide a film, barrier, coating or as a gelling agent.
25. Use of starch according to claim 23, to prepare resistant starch compositions.
26. Use of starch according to any one of claims 1-22 in the preparation or processing of corrugating adhesives, biodegradable products, packaging, glass fibers and textiles.
27. A nucleotide sequence encoding an effective portion of a class A starch branching

enzyme (SBE) obtainable from potato plants.

28. A nucleotide sequence according to claim 27, encoding a polypeptide comprising substantially the amino acid sequence of residues 49 to 882 of the sequence shown in Figure 5.

29. A nucleotide sequence according to claim 27 or 28, comprising substantially the sequence of nucleotides 289 to 2790 of the sequence shown in Figure 5, or a functional equivalent thereof.

30. A nucleotide sequence according to claim 29, further comprising the sequence of nucleotides 145 to 288 of the sequence shown in Figure 5, or a functional equivalent thereof.

31. A nucleotide sequence according to claim 27, comprising the sequence of nucleotides 228 to 2855 of the sequence labelled psbe2con.seq in Figure 8, or a functional equivalent thereof.

32. A nucleotide sequence according to claim 27, comprising the sequence of nucleotides 57 to 2564 of the sequence labelled as psbe2con.seq in Figure 12, or a functional equivalent thereof.

33. A nucleotide sequence according to any one of claims 27 to 32, comprising an in-frame ATG start codon, and optionally including a 5' and/or a 3' untranslated region.

34. A nucleotide sequence according to claim 27, comprising the sequence of nucleotides 45 to 3200 of the sequence labelled as psbe2con.seq in Figure 8, or a functional equivalent thereof.

35. A nucleic acid construct comprising a sequence in accordance with any one of claims 27 to 34.

36. An expression vector comprising a nucleic acid construct according to claim 35.
37. A host cell into which has been introduced a sequence in accordance with any one of claims 27 to 36.
38. An effective portion of a class A SBE polypeptide obtainable from potato plants and encoded by a nucleotide sequence in accordance with any one of claims 27 to 36.
39. A polypeptide according to claim 38, comprising substantially the sequence of amino acids 49 to 882 of the sequence shown in Figure 5, or a functional equivalent thereof.
40. A polypeptide according to claim 38 or 39, comprising the sequence of amino acids 1 to 48 of the sequence shown in Figure 5.
41. A polypeptide in accordance with any one of claims 38, 39 or 40 in substantial isolation from other plant-derived constituents.
42. A method of altering the characteristics of a plant, comprising introducing into the plant a portion of a nucleotide sequence in accordance with any one of claims 27 to 36, operably linked to a suitable promoter active in the plant, so as to affect the expression of a gene present in the plant.
43. A method according to claim 42, wherein the nucleotide sequence is operably linked in the anti-sense orientation to a suitable promoter active in the plant.
44. A method according to claim 42, wherein the introduced sequence comprises one or more of the following operably linked in the sense orientation to a promoter active in the plant, so as to cause sense suppression of an enzyme naturally expressed in the plant: a 5' untranslated region, a 3' untranslated region, or a coding region of the potato SBE class A starch branching enzyme.
45. A method according to any one of claims 42, 43 or 44, further comprising

introducing into the plant one or more further sequences.

46. A method according to claim 45, wherein one or more of the further sequences are operably linked in the anti-sense orientation to a suitable promoter active in the plant.

47. A method according to claim 45 or 46, wherein the further sequence comprises a portion of a class B SBE nucleotide sequence.

48. A method according to any one of claims 42 to 47, effective in altering the starch composition of a plant.

49. A plant or plant cell having characteristics altered by the method of any one of claims 42 to 48, or the progeny of such a plant, or part of such a plant.

50. A plant according to claim 49, selected from one of the following: potato, pea, tomato, maize, wheat, rice, barley, sweet potato, and cassava.

51. A tuber or other storage organ from a plant according to claim 49 or 50.

52. Use of a tuber or other storage organ according to claim 51, in the preparation and/or processing of a foodstuff.

53. A plant according to claim 49 or 50, containing starch which, as extracted from the plant by wet milling at ambient temperature, has an elevated viscosity onset temperature as judged by viscoamylograph conducted according to the protocol defined in claim 7, compared to starch extracted from a similar, but unaltered, plant.

54. A plant according to claim 53, wherein the viscosity onset temperature is elevated by an amount in the range of 10 to 25°C.

55. A plant according to claim 49 or 50, containing starch which, as extracted from the plant by wet milling at ambient temperature, has a decreased peak viscosity as judged by

viscoamylograph conducted according to the protocol defined in claim 7, compared to starch extracted from a similar, but unaltered, plant.

56. A plant according to claim 55, wherein the peak viscosity is decreased by an amount in the range of 240 to 700 SNU's.

57. A plant according to claim 49 or 50, containing starch which, as extracted from the plant by wet milling at ambient temperature, has an increased pasting viscosity as judged by viscoamylograph conducted according to the protocol defined in claim 7, compared to starch extracted from a similar, but unaltered, plant.

58. A plant according to claim 57, wherein the pasting viscosity is increased by an amount in the range of 37 to 260 SNU's.

59. A plant according to claim 49 or 50, containing starch which, as extracted from the plant by wet milling at ambient temperature, has an increased set-back viscosity as judged by viscoamylograph conducted according to the protocol defined in claim 7, compared to starch extracted from a similar, but unaltered, plant.

60. A plant according to claim 59, wherein the set-back viscosity is increased by an amount in the range of 224 to 313 SNU's.

61. A plant according to claim 49 or 50, containing starch which, as extracted from the plant by wet milling at ambient temperature, has a decreased set-back viscosity as judged by viscoamylograph conducted according to the protocol defined in claim 7, compared to starch extracted from a similar, but unaltered, plant.

62. A plant according to claim 49 or 50, containing starch which, as extracted from the plant by wet milling at ambient temperature, has an elevated apparent amylose content as judged by iodometric assay according to the method of Morrison & Laignelet, compared to starch extracted from a similar, but unaltered, plant.

63. A plant according to claim 49 or 50, containing starch which was extracted from the plant, has a phosphorus content in excess of 200mg/100grams dry weight starch.
64. Starch obtainable from a plant according to any one of claims 49, 50 or 53 - 63.
65. Starch according to claim 64 and further in accordance with any one of claims 1 - 22.
66. A method of modifying starch *in vitro*, comprising treating starch under suitable conditions with an effective amount of a polypeptide in accordance with any one of claims 38 to 41.
67. A potato plant or part thereof which, in its wild type possesses an effective SBE A gene, but which plant has been altered such that there is no effective expression of an SBE A polypeptide within the cells of at least part of the plant.
68. A potato plant according to claim 67, wherein the alteration is effected by a method according to any one of claims 42-48.

1/75

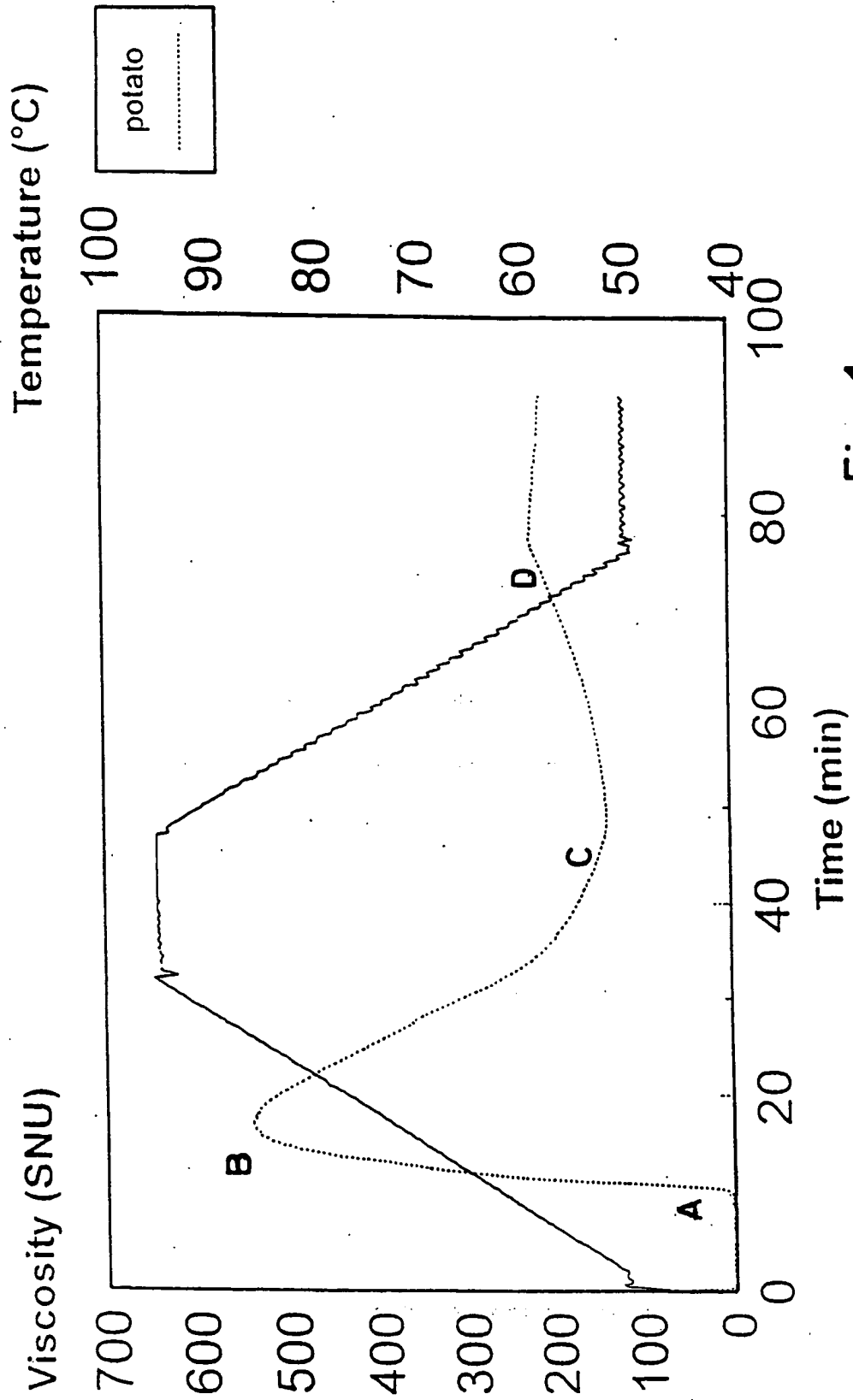


Fig. 1

2/75

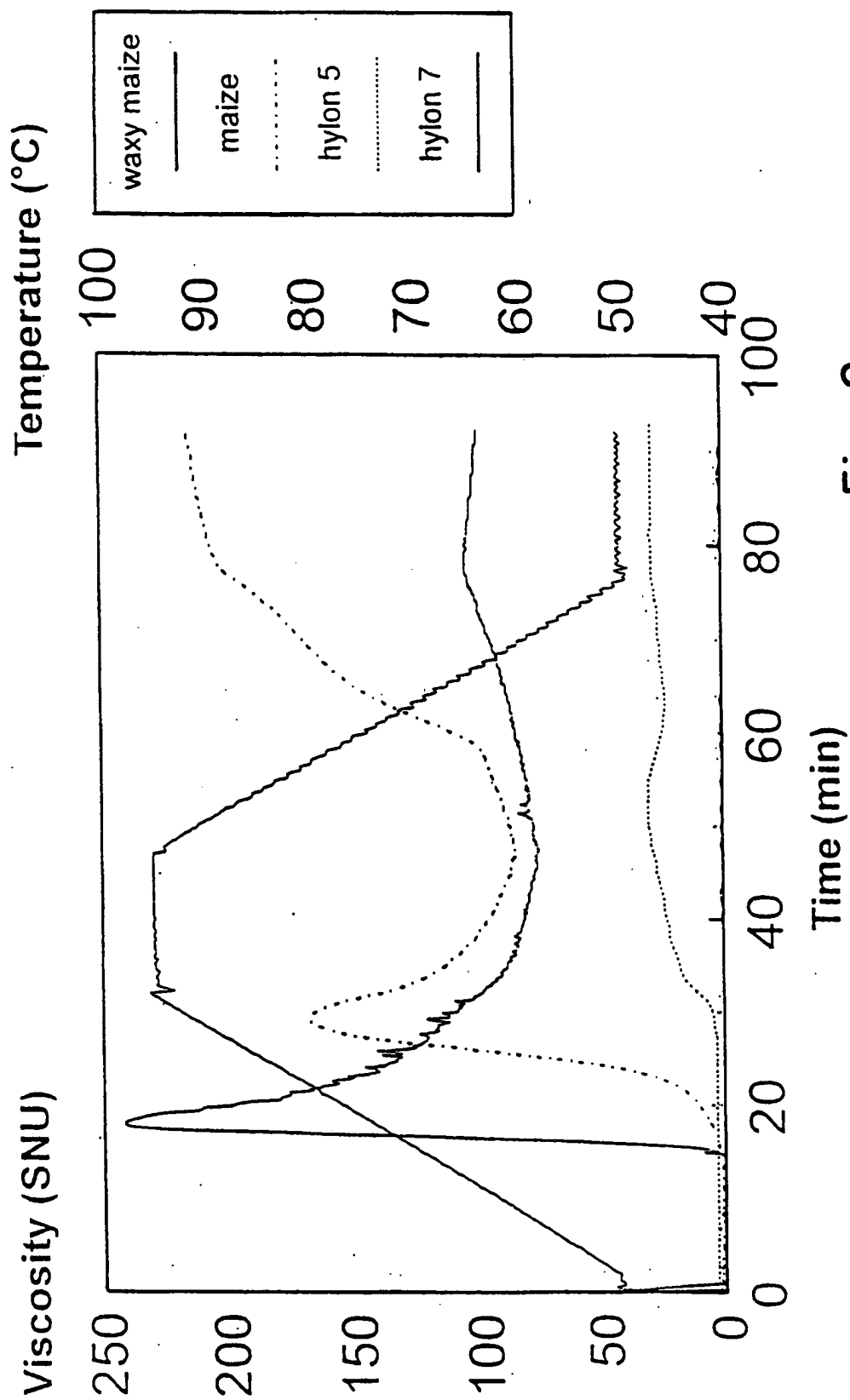


Fig. 2

3/75

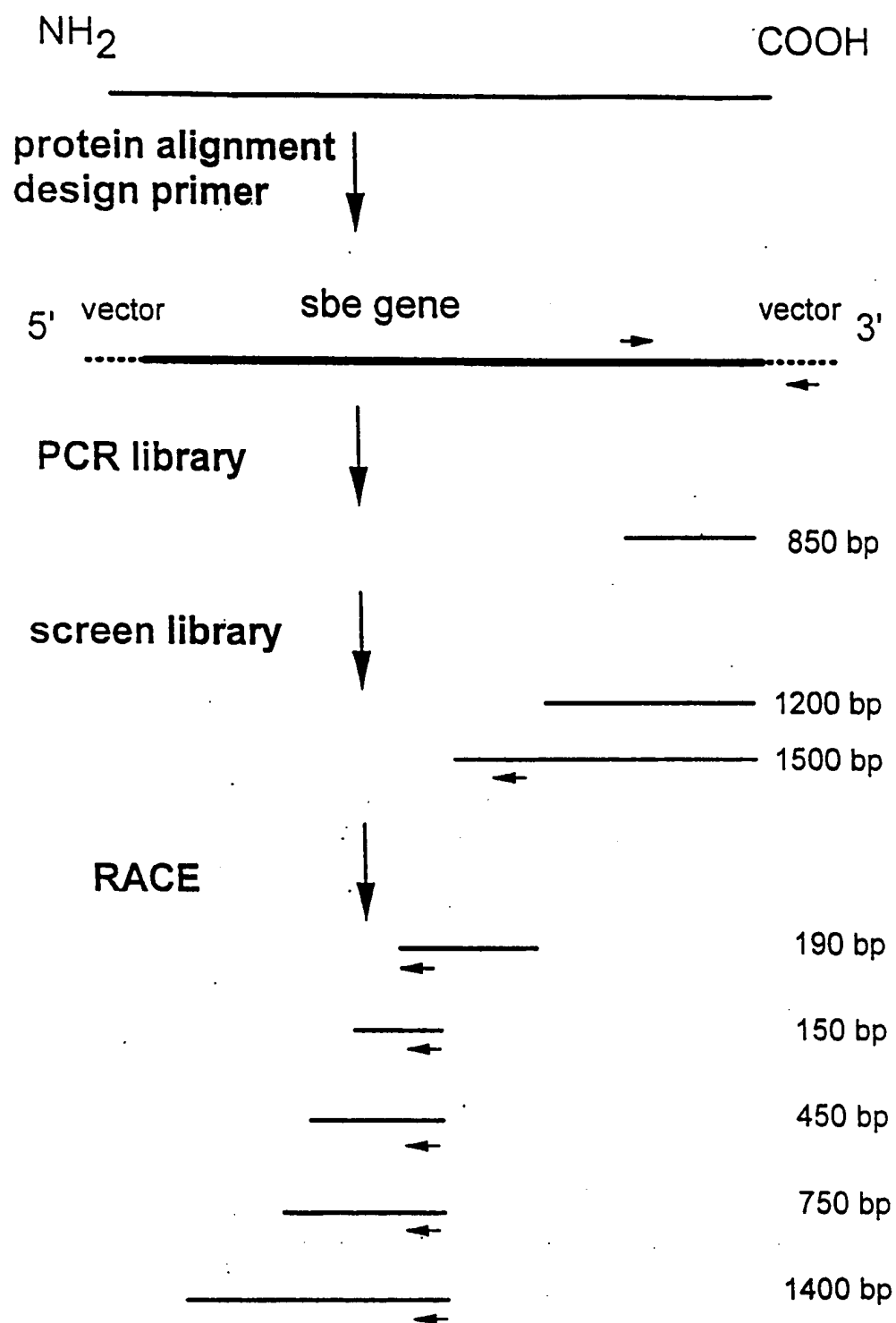


Fig. 3

4/75
Fig.4a
Sheet 2

Majority	P	A	S	P	T	I	D	R	G	I	A	L	H	K	M	I	H	L	I	T	M	G	L	G	G	E	G	Y	L	N	F	M	G	N	
maize 2	P	S	T	P	T	I	D	R	G	I	A	L	H	K	M	I	R	L	I	T	M	G	L	G	G	E	G	Y	L	N	F	M	G	N	
pea 1	P	S	T	P	L	I	D	R	G	I	A	L	H	K	M	I	R	L	I	T	M	G	L	G	G	E	G	Y	L	N	F	M	G	N	
maize 1	P	A	S	P	T	I	D	R	G	I	A	L	H	K	M	I	H	F	I	T	M	A	L	G	G	E	G	Y	L	N	F	M	G	N	
rice 1	P	A	S	P	T	I	N	R	G	I	A	L	H	K	M	I	H	F	I	T	M	A	L	G	G	E	G	Y	L	N	F	M	G	N	
potato1	D	A	S	P	V	V	D	A	G	I	A	L	H	K	M	I	H	F	I	T	M	A	L	G	G	E	G	Y	L	N	F	M	G	N	
human	P	F	T	P	V	I	D	R	G	I	A	L	H	K	M	I	R	L	I	T	H	G	L	G	G	E	G	Y	L	N	F	M	G	N	
Majority	F	S	L	G	D	A	D	H	L	R	Y	K	G	M	N	A	F	D	Q	A	M	N	A	L	E	E	K	F	S	F	L	A	S	S	
maize 2	F	D	L	G	D	A	D	Y	L	R	Y	H	G	M	Q	E	F	D	Q	A	M	Q	H	L	E	Q	K	Y	E	F	M	T	S	E	
pea 1	F	D	L	G	D	A	D	Y	L	R	Y	H	G	M	Q	E	F	D	Q	A	M	Q	H	L	E	Q	K	Y	E	F	M	T	S	E	
maize 1	W	S	L	V	D	T	T	D	H	L	R	Y	K	Y	M	N	A	F	D	Q	A	M	N	A	L	E	E	R	F	S	F	L	S	S	
rice 1	W	S	L	V	D	T	T	D	H	L	R	Y	K	Y	M	N	A	F	D	Q	A	M	N	A	L	E	E	R	F	S	F	L	S	S	
potato1	W	N	L	A	D	S	E	H	L	R	Y	K	F	F	L	N	A	F	D	R	D	R	L	E	E	R	Y	G	W	L	A	A	P		
human	F	H	L	T	D	D	D	L	L	R	Y	K	F	F	L	N	A	F	D	R	D	R	L	E	E	R	Y	G	W	L	A	A	P		
Majority	K	V	G	C	D	L	P	G	K	Y	K	V	A	L	D	S	D	A	L	V	F	G	G	H	G	R	V	G	H	D	V	D	H	F	
maize 2	R	I	G	C	R	K	P	G	V	Y	K	V	V	L	D	S	D	A	G	L	F	G	G	F	S	R	I	H	H	A	A	E	H	F	
pea 1	K	V	G	C	L	K	P	G	K	Y	K	I	V	L	D	S	D	A	D	T	L	F	G	G	F	N	R	L	N	H	T	A	E	H	F
maize 1	K	V	G	C	D	L	P	G	K	Y	R	V	A	L	D	S	D	A	L	V	F	G	G	H	G	R	V	G	H	D	V	D	H	F	
rice 1	K	V	G	C	D	L	P	G	K	Y	R	V	A	L	D	S	D	A	L	V	F	G	G	H	G	R	V	G	H	D	V	D	H	F	
potato1	K	V	G	C	D	L	P	G	K	Y	R	V	A	L	D	S	D	A	W	E	F	G	G	H	G	R	A	G	H	D	V	D	H	F	
human	R	V	G	T	A	L	P	G	K	F	K	I	V	L	D	S	D	A	A	E	Y	G	G	H	Q	R	L	D	H	S	T	D	F	F	

Fig. 4a SHEET 1

5/75

E F G H P E W I D F P R E - - - - - G N N W S Y D K C R R O	
E F G H P E W I D F P R R G P Q R L P S G K F I P	R 666
E F G H P E W I D F P R G E Q H L P N G K I V	R 713
E F G H P E W I D F P R E - - - - -	Q 624
E F G H P E W I D F P R E - - - - -	Q 618
E F G H P E W I D F P R S E - - - - -	Q 638
E F G H P E W L D F P R K - - - - -	Q 566
K Q I V S D K N E G D K V I V F E R G D L V F V F N F H P N N S Y E G Y	
H Q Y I S R K H E E D K V I V F F E K G D L V F V F F N F H C N N S Y F D Y	736
H Q Y I S R K N E G D R V I I F F E R G D L V F V F F N F H W T N S Y S D Y	783
K Q I V S D M N D E K V I V F F E R G D L V F V F F N F H P K T Y E G Y	694
K Q I V S D M N E K V I V F F E R G D L V F V F F N F H P K T Y K G Y	688
K Q I V S S M D D N K V I V F F E R G D L V F V F F N F H P N T Y E G Y	708
Q A Y V S S E K H E G N K I I A F F E R A G L L F I F F N F H P S K S Y T D Y	636
T S P E G - P G V P E T N F N N R P N S F K V L S P S R T C V A Y Y R V	
T A - - - - - D C S H D N R P P Y S F S V Y T P S R T C V V Y A P V	798
T S - - - - - E G W Y D D R P P R S F L V V Y A P S R T A V V Y A L A	845
T S P E G V P G V P E T N F N N R P P N S F F K V L S P P P R T C V A Y Y R V	764
T S P E G M P G V P E T N F N N R P P N S F F K V L S P P A R T C V A Y Y R V	758
S E - - - - - A F E H N G R P Y S L L V Y I P S R V A L I L Q N V	778 698

Fig. 4a SHEET 2

6/75

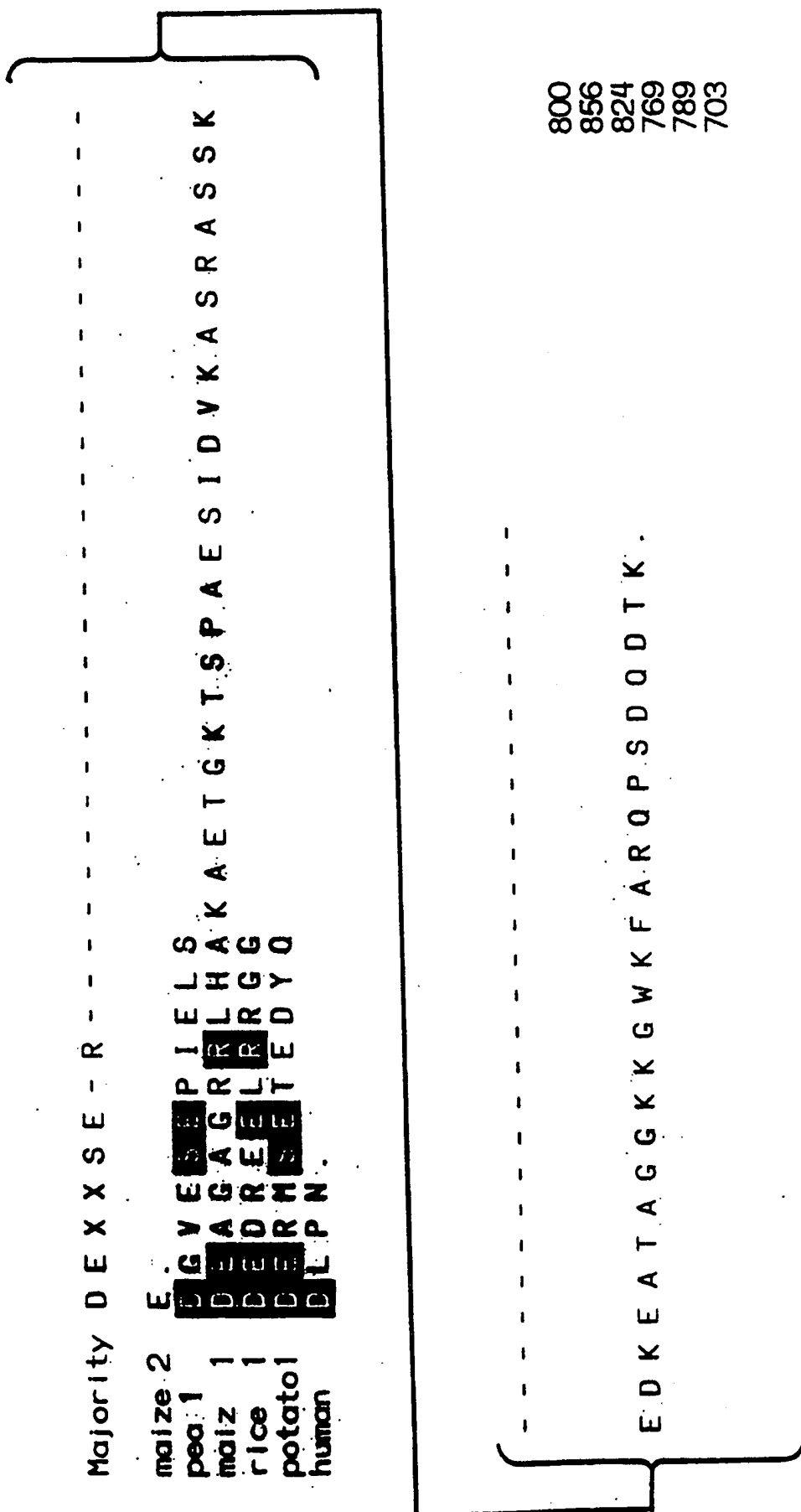
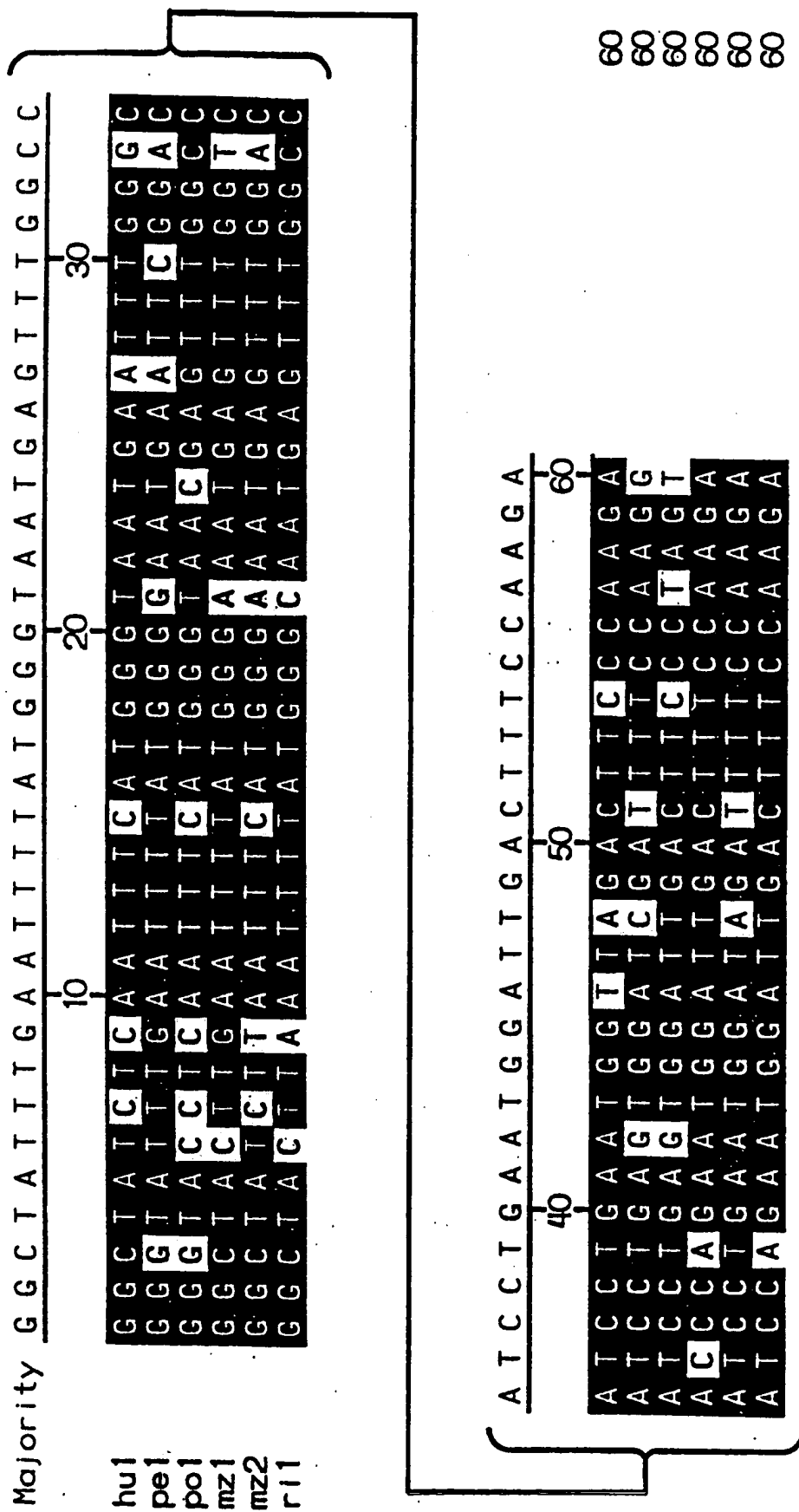


Fig. 4a SHEET 3

7175



8/75

TTGATGGGGCCTTGAAGCTCAGCAATTTGACACTCAGTTAGTTACA
+ + + + +
AACTACCCCGGAAGCTTGAGTCGTTAAACTGTGAGTCAATCAATGT

AAGGAATGAATAAAAGGATAGATTTGTAAAAACCCTAAGGAGAGA
+ + + + +
TTCCTTACTTATTTTCCTATCTAAACATTTTGGGATTCCTCTCT

M N K R I D L

GTTCCATCAGTGTACAAATCTAATGGATTCAGCAGTAATGGTGAT
+ + + + +
CAAGGTAGTCACATGTTTAGATTACCTAAGTCGTCATTACCACTA
V P S V Y K S N G F S S N G D

Bgl II

EcoR I

TCACGGAAGATCTTGGCTGAAAAGTCTTCTTACAATTCCGAATTC
+ + + + +
AGTGCCTTCTAGAACCGACTTTTCAGAAGAATGTTAAGGCTTAAG
S R K I L A E K S S Y N S E F

ACCCAGAGTGATAGCTCCTCATCCTCAACAGACCAATTTGAGTTC
+ + + + +
TGGGTCTCACTATCGAGGAGTAGGAGTTGTCTGGTTAAACTCAAG
T Q S D S S S S S T D Q F E F

AGTTCAACAATGGAACACGCTAGCCAGATTAAAGTGAAGAAGAT
+ + + + +
TCAAGTTGTTACCTTGTGCGATCGGTCTAATTTTGACTCTTGCTA
S S T M E H A S Q I K T E N D

GATTTTGCTTCATCACTACAAGTACAAGAAGGTGGTAAACTGGAG
+ + + + +
CTAAAACGAAGTAGTGATGTTGATGTTCTTCCACCATTTGACCTC
D F A S S L Q L Q E G G K L E

Fig 5
Sheet 2

Fig. 5 SHEET 1

Bgl II

90

180

270

360

450

540

630

10/75

TCTGATAGGATCAGAGAGAGGGGCATCCCTCCACCTGGACTTGGT
AGACTATCCTAGTCTCTCTCCCCGTAGGGAGGTGGACCTGAACCA
S D R I R E R G I P P P G L G

CACCTTGATTACAGGTATTACAGTACAAGAACTGAGGGAGGCA
GTGGAACTAATGTCCATAAGTGTCATGTTCTTTGACTCCCTCCGT
H L D Y R Y S Q Y K K L R E A

GAAAAAATGGGTTTCACTCGTAGTGCTACAGGTATCACTTACCGT
CTTTTTTACCCAAAGTGAGCATCACGATGTCCATAGTGAATGGCA
E K M G F T R S A T G I T Y R

AACAATTGGGACGCAAATGCTGACATTATGACTCGGAATGAATTT
TTGTTAACCCTGCGTTTACGACTGTAATACTGAGCCTTACTTAAA
N N W D A N A D I M T R N E F

GCAATTCCTCATGGGTCCAGAGTGAAGATACGTATGGACACTCCA
CGTTAAGGAGTACCCAGGTCTCACTTCTATGCATACCTGTGAGGT
A I P H G S R V K I R M D T P

Fig.5
Sheet4

Fig. 5 SHEET 3

11/75

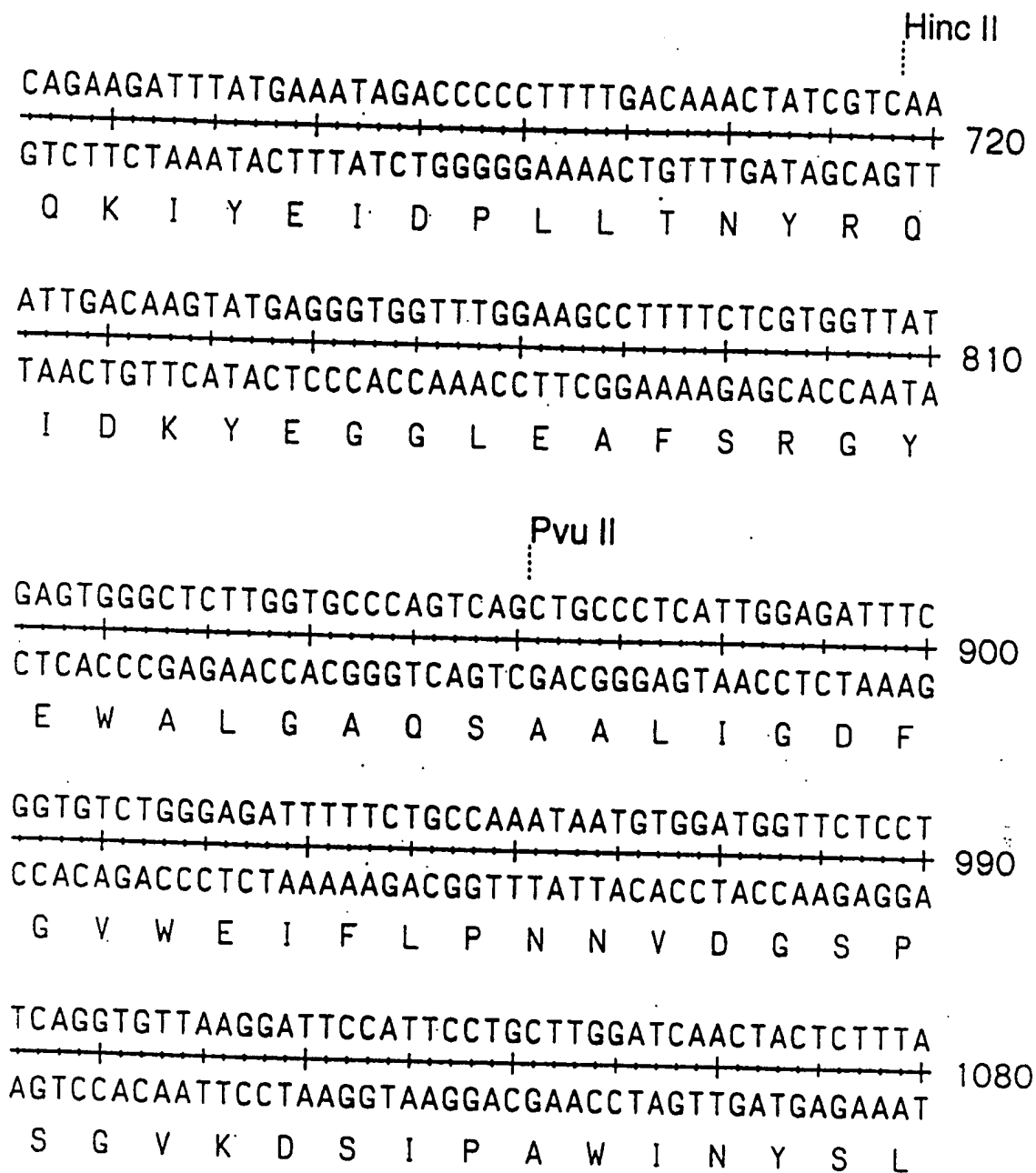


Fig. 5 SHEET 4

12/75

CAGCTTCCTGATGAAATTCCATATAATGGAATACATTATGATCCA
GTCGAAGGACTACTTTAAGGTATATTACCTTATGTAATACTAGGT
Q L P D E I P Y N G I H Y D P

CCAAAGTCGCTGAGAATATATGAATCTCATATTGGAATGAGTAGT
GGTTTCAGCGACTCTTATATACTTAGAGTATAACCTTACTCATCA
P K S L R I Y E S H I G M S S

HinD III

CTTCCTCGCATAAAAAAGCTTGGGTACAATGCGCTGCAAATTATG
GAAGGAGCGTATTTTTTTCGAACCCATGTTACGCGACGTTTAATAC
L P R I K K L G Y N A L Q I M

ACAAATTTTTTGCACCAAGCAGCCGTTTTTGAACGCCCGACGAC
TGTTTAAAAAACGTGGTTCGTGCGCAAAACCTTGCGGGCTGCTG
T N F F A P S S R F G T P D D

CTCATGGACATTGTTACAGCCATGCATCAAATAATACTTTAGAT
GAGTACCTGTAACAAGTGTCGGTACGTAGTTTATTATGAAATCTA
L M D I V H S H A S N N T L D

Fig.5
Sheet
6

Fig. 5 SHEET 5

13/75

CCCGAAGAGGAGAGGTATATCTTCCAACACCCACGGCCAAAGAAA 1170
GGGCTTCTCCTCTCCATATAGAAGGTTGTGGGTGCCGGTTTCTTT
P E E E R Y I F Q H P R P K K

Xmn I
CCGGAGCCTAAAATTAAC TCATACGTGAATTTTAGAGATGAAGTT 1260
GGCCTCGGATTTTAATTGAGTATGCACTTAAAATCTCTACTTCAA
P E P K I N S Y V N F R D E V

GCTATTCAAGAGCATTCTTATTACGCTAGTTTTGGTTATCATGTC 1350
CGATAAGTTCTCGTAAGAATAATGCGATCAAAACCAATAGTACAG
A I Q E H S Y Y A S F G Y H V

CTTAAGTCTTTGATTGATAAAGCTCATGAGCTAGGAATTGTTGTT 1440
GAATTCAGAACTAACTATTTGAGTACTCGATCCTTAACAACAA
L K S L I D K A H E L G I V V

GGACTGAACATGTTTGACTGCACCGATAGTTGTTACTTTCACTCT 1530
CCTGACTTGTACAACTGACGTGGCTATCAACAATGAAAGTGAGA
G L N M F D C T D S C Y F H S

Fig. 5 SHEET 6

14/75

Sac I

GGAGCTCGTGGTTATCATTGGATGTGGGATTCCCGCCTCTTTAAC
CCTCGAGCACCAATAGTAACCTACACCCTAAGGGCGGAGAAATTG
G A R G Y H W M W D S R L F N

TGGTGGTTGGATGCGTTCAAATTTGATGGATTTAGATTTGATGGT
ACCACCAACCTACGCAAGTTTAAACTACCTAAATCTAAACTACCA
W W L D A F K F D G F R F D G

ACTGGGAACCTACGAGGAATACTTTGGACTCGCAACTGATGTGGAT
TGACCCTTGATGCTCCTTATGAAACCTGAGCGTTGACTACACCTA
T G N Y E E Y F G L A T D V D

TTCCCAGATGCAATTACCATTGGTGAAGATGTTAGCGGAATGCCG
AAGGGTCTACGTTAATGGTAACCACTTCTACAATCGCCTTACGGC
F P D A I T I G E D V S G M P

CGGCTGCATATGGCAATTGCTGATAAACGGATTGAGTTGCTCAAG
GCCGACGTATACCGTTAACGACTATTTGCCTAACTCAACGAGTTC
R L H M A I A D K R I E L L K

ACAAATAGAAGATGGTCGGAAAAGTGTGTTTTCATACGCTGAAAGT
TGTTTATCTTCTACCAGCCTTTTCACACAAAGTATGCGACTTTCA
T N R R W S E K C V S Y A E S

Fig 5
Sheet 8

Fig. 5 SHEET 7

15/75

TATGGAACTGGGAGGTACTTAGGTATCTTCTCTCAAATGCGAGA
+-----+
ATACCTTTGACCCTCCATGAATCCATAGAAGAGAGTTTACGCTCT 1620
Y G N W E V L R Y L L S N A R

GTGACATCAATGATGTATATTCACCACGGATTATCGGTGGGATTCT
+-----+
CACTGTAGTTACTACATATAAGTGGTGCCTAATAGCCACCCTAAG 1710
V T S M M Y I H H G L S V G F

Hinc II
|

GCTGTTGTGTATCTGATGCTGGTCAACGATCTTATTCATGGGCTT
+-----+
CGACAACACATAGACTACGACCAGTTGCTAGAATAAGTACCCGAA 1800
A V V Y L M L V N D L I H G L

ACATTTTGTATTCCCGTCCAAGAGGGGGGTGTTGGCTTTGACTAT
+-----+
TGTA AACATAAGGGCAGGTTCTCCCCCACAACCGAAACTGATA 1890
T F C I P V Q E G G V G F D Y

AAACGGGATGAGGATTGGAGAGTGGGTGATATTGTTTCATACACTG
+-----+
TTTGCCCTACTCCTAACCTCTCACCCTACTATAACAAGTATGTGAC 1980
K R D E D W R V G D I V H T L

CATGATCAAGCTCTAGTCGGTGATAAACTATAGCATTCTGGCTG
+-----+
GTA TAGTTTCGAGATCAGCCACTATTTTGATATCGTAAGACCGAC 2070
H D Q A L V G D K T I A F W L

Fig. 5 SHEET 8

16/75

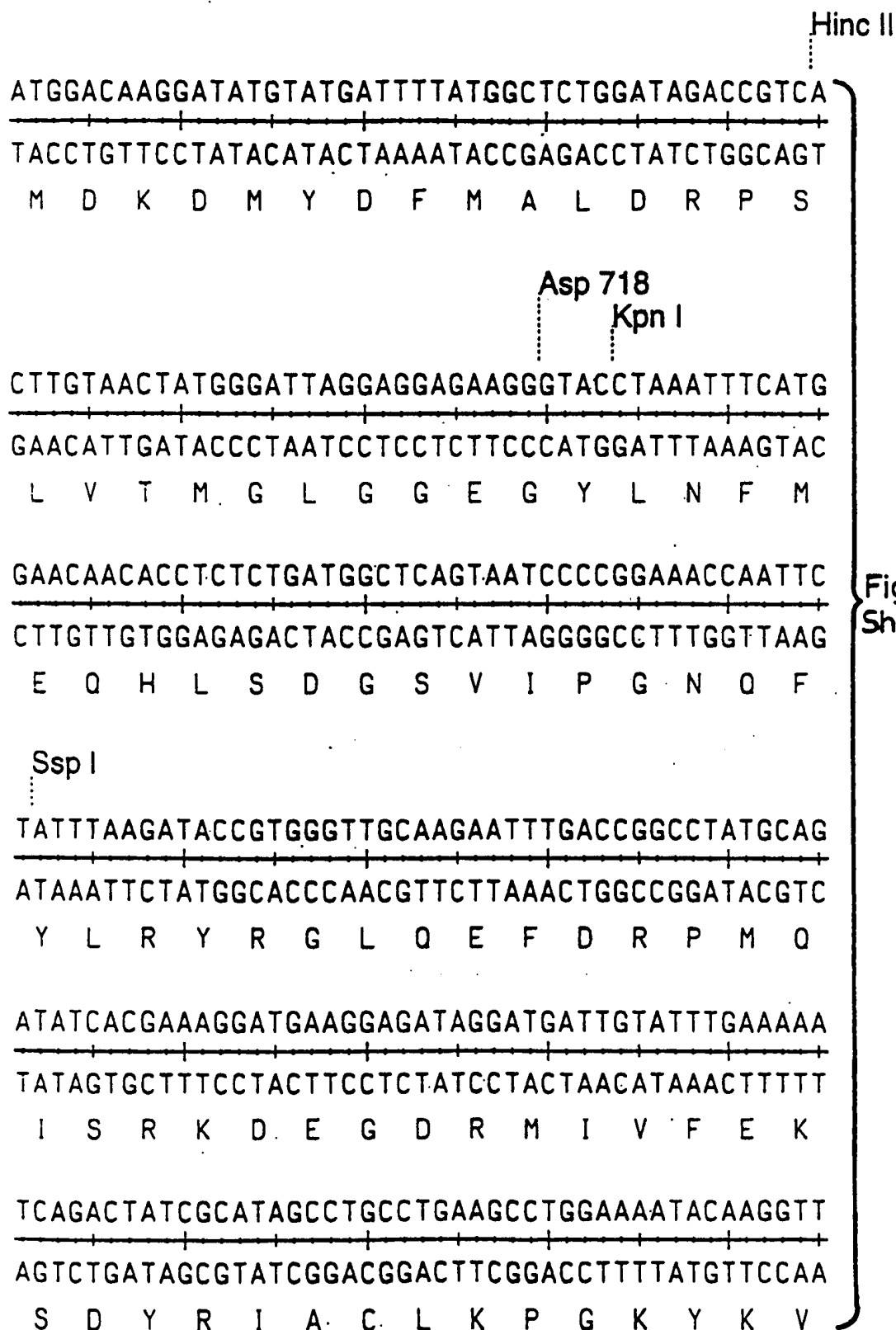
Fig.5
Sheet10

Fig.5 SHEET 9

SUBSTITUTE SHEET (RULE 26)

17/75

ACATCATTAAATAGATCGTGGGATAGCATTGCACAAGATGATTAGG
TGTAGTAATTATCTAGCACCTATCGTAACGTGTTCTACTAATCC
T S L I D R G I A L H K M I R

EcoR I

GGAAATGAATTCGGCCACCCTGAGTGGATTGATTTCCCTAGGGCT
CCTTTACTTAAGCCGGTGGGACTCACCTAACTAAAGGGATCCCGA
G N E F G H P E W I D F P R A

AGTTATGATAAATGCAGACGGAGATTTGACCTGGGAGATGCAGAA
TCAATACTATTTACGTCTGCCTCTAACTGGACCCTCTACGTCTT
S Y D K C R R R F D L G D A E

TATCTTGAAGATAAATATGAGTTTATGACTTCAGAACACCAGTTC
ATAGAACTTCTATTTATACTCAAATACTGAAGTCTTGTGGTCAAG
Y L E D K Y E F M T S E H Q F

GGAAACCTAGTTTTTGTCTTTAATTTTCACTGGACAAAAAGCTAT
CCTTTGGATCAAAAACAGAAATTAAGTACCTGTTTTTTCGATA
G N L V F V F N F H W T K S Y

GCCTTGGACTCAGATGATCCACTTTTTTGGTGGCTTCGGGAGAATT
CGGAACCTGAGTCTACTAGGTGAAAAACCACCGAAGCCCTCTTAA
A L D S D D P L F G G F G R I

Fig. 5 SHEET 10

18/75

Ssp I

GATCATAATGCCGAATATTTACCTTTGAAGGATGGTATGATGAT

CTAGTATTACGGCTTATAAAGTGGAACCTTCCTACCATACTACTA

D H N A E Y F T F E G W Y D D

GTCTATGCACTAGTAGACAAAGAAGAAGAAGAAGAAGAAGAA

CAGATACGTGATCATCTGTTTCTTCTTCTTCTTCTTCTTCTTCTT

V Y A L V D K E E E E E E E E

TGAACGAACTTGTGATCGCGTTGAAAGATTTGAACGCTACATAGA

ACTTGCTTGAACACTAGCGCAACTTTCTAACTTGCGATGTATCT

Fig 5
Sheet
12

TCATGTGACACAAGGTTTGCAATTCTTTCCACTATTAGTAGTGCA

AGTACACTGTGTTCCAAACGTTAAGAAAGGTGATAATCATCACGT

EcoR I

Pst I

GATGAATTTATGTCGAATGCTGGGACGATCGAATTCCTGCAGGCC

CTACTTAAATACAGCTTACGACCCTGCTAGCTTAAGGACGTCCGG

Fig. 5 SHEET 11

19/75

CGTCCTCGTTCAATTATGGTGTATGCACCTTGTAACAGCAGTG 2700
GCAGGAGCAAGTTAATACCACATACGTGGAACATTTTGTCTGTCAC
R P R S I M V Y A P C K T A V

GAAGAAGAAGTAGCAGCAGTAGAAGAAGTAGTAGTAGAAGAAGAA 2790
CTTCTTCTTCATCGTCGTCATCTTCTTCATCATCATCTTCTTCTT
E E E V A A V E E V V V E E E

Ssp I

GCTTCTTGACGTATCTGGCAATATTGCATCAGTCTTGGCGGAATT 2880
CGAAGAACTGCATAGACCGTTATAACGTAGTCAGAACCGCCTTAA

Cla I

ACGATATACGCAGAGATGAAGTGCTGAACAAACATATGTAAATC 2970
TGCTATATGCGTCTCTACTTCACGACTTGTTTGTATACATTTTAG

GGGGGACCCCTTAGTTCT 3033
CCCCCTGGGGAATCAAGA

Fig. 5 SHEET 12

20/75

✓180 ✓190 ✓200 ✓210 ✓220
 IYEIDPLLTNYRQHLDRYSQYKKLREAI DKYEGGLEAFSRGYEKMGTFR
 : : : DP L. Y : H: . R : : Y . : I: KYEG LE. F: : GY K. GF. R
 LLNLDPTLEPYLDHFRHRMKRYVDQKMLIEKYEGPLEEFAQGYLKFGFNRR
 ^100 ^110 ^120 ^130 ^140
 ✓230 ✓240 ✓250 ✓260 ✓270
 SATGITYREWALGAQSAALIGDFNNWDANADIMTRNEFGVWEIFLPNNVD
 ... I. YREWA : AQ. A. : IGDFN. W: : : : : M. : : : FGVW. I : P: VD
 EDGCIVYREWAPAAQEA EVIGDFNGWNGSNHMMEKDQFGVWSIRIPD-VD
 ^150 ^160 ^170 ^180 ^190
 ✓280 ✓290 ✓300 ✓310 ✓320
 GSPAIPHGSRVKIRMDTPSGV-KDSIPAWINYSLQLPDEI--PYNGIHYD
 : . P. IPH. SRVK: R. : : GV D. IPAWI: Y: . : : : PY: G: . D
 SKPVIPHNSRVKFRFKHGNVWVDRIPAWIKYATADATKFAAPYDGVYWD
 ^200 ^210 ^220 ^230 ^240
 ✓330 ✓340 ✓350 ✓360 ✓370
 PPEEERYIFQHPRPKPKSLRIYESHIGMSSPEPKINSYVNFRRDEVLPRI
 PP . ERY F: . PRP KP: : RIYE: H: GMSS: EP: : NSY : F D: VLPRI
 PPPSERYHFKYPRPPKPRAPRIYEAHVGMSSSEPRVNSYREFADDVLPRI
 ^250 ^260 ^270 ^280 ^290
 ✓380 ✓390 ✓400 ✓410 ✓420
 KKLGYNALQIMAIQEHSYYASFGYHVTNFFAPSSRFGTPDDLKSLIDKAH
 K . YN: : Q: MAI EHSYY: SFGYHVTNFFA S: R: G. P: DLK LIDKAH
 KANNYNTVOLMAIMEHSYYGSFGYHVTNFFAVSNRYGNPEDLKYLIDKAH
 ^300 ^310 ^320 ^330 ^340
 ✓430 ✓440 ✓450 ✓460 ✓470
 ELGIVVLMDIVHSHASNNTLDGLNMFDC---TDSCYFHSGARGYHWMWDS
 . LG: VL: D: VHSHASN. DGLN FD : : : YFH: G. RGYH : WDS
 SLGLQVLVDVHSHASNNTDGLNGFDIGQGSQESYFHAGERGYHKLWDS
 ^350 ^360 ^370 ^380 ^390
 ✓480 ✓490 ✓500 ✓510 ✓520
 RLFNYGNWEVRLRYLLSNARWWLDAFKFDGFRFDGVTSMYIHHGLSVGFT
 RLFNY: NWEVLR: LLSN RWL: : : : FDGFRFDG: TSM: Y: HHG: : : GFT
 RLFNYANWEVLRFLLSNLRWWLEEYNFDGFRFDGITSMLYVHHGINMGFT
 ^400 ^410 ^420 ^430 ^440
 ✓530 ✓540 ✓550 ✓560 ✓570
 GNYEEYFGLATDVAVVYLMVNDLIHGLFPDAITIGEDVSGMPTFCIPV
 GNY: EYF: ATDVAVVYLM. N: LIH : FPDA. . I: EDVSGMP. : . PV
 GNYNEYFSEATDVAVVYLMANLIHKIFPDATVIAEDVSGMPGLSRPV
 ^450 ^460 ^470 ^480 ^490
 ✓580 ✓590 ✓600 ✓610 ✓620
 OEGGVGFDYRLHMAIADKRIELLK-KRDEDWRVGDIVHTLTNRRWSEKCV
 EGG: GFDYRL MAI: DK: I: LK K. DEDW. : : : : LTNRR. : EKC:
 SEGGIGFDYRLAMAI PKWIDYLNKNKNDWDMSMKEVTSSLTNRRYTEKC I
 ^500 ^510 ^520 ^530 ^540

Fig. 6 SHEET 1

21/75

630 640 650 660 670
 SYAESHDOALVGDKTIAFWLMDKDMYDFMALDRPSTSLIDRGIALHKMIR
 : YAESHDOQ: : VGDKTIAF LMDK: MY. M: : : : : DRGIALHKMI:
 AYAESHDOQSIVGDKTIAFLLMDKEMYSGMSCLTDASPVVDRGIALHKMIH
 550 560 570 580 590
 680 690 700 710 720
 LVTMGLGGEGYLNFMGNEFGHPEWIDFPRAEQHLSGSGVIPGNQFSYDKC
 : TM: LGGEGYLNFMGNEFGHPEWIDFPR GN: . SYDKC
 FFTMALGGEGYLNFMGNEFGHPEWIDFPR-----EGNNWSYDKC
 600 610 620 630
 730 740 750 760 770
 RRRFDLGDAEYLRGRGLQEFDRPMQYLEDKYEFTSEHQFISRKDEGDRM
 RR: . : L: D: E. LRY: : : . FDR: M: L: : K: . F: : S: . Q: : S: . D: : : : :
 RRQWNLDSEHLRYKFMNAFDRAMNSLDEKFSFLASGKQIVSSMDDDNKV
 640 650 660 670 680
 780 790 800 810 820
 IVFEKGNLVFVFNFWTKSYSDYRIACLKPGKYKVALDSDDPLFGGFGRI
 : VFE: G: LVFVFNFH . : : Y: : Y: : : C PGKY: VAL: SD. FGG GR
 VVFERGDLVFVFNFHHPNNTYEGYKVGCDLPGKYRVALGSDAWEFGGHGRA
 690 700 710 720 730
 830 840 850 860
 DHNAEYFT-----FEGWYDDRPRSIMVYAPCKTAVVYALVDKEEEEE
 : H: . : . FT E. : : : RP. S: . V : P : T V. Y VD. . E.
 GHDVDHFTSPEGIPGVPETNFNGRPNFSKVLSPARTCVAYYRVDERMSET
 740 750 760 770 780
 870
 EEEEEEV
 E: . : : :
 EDYQTDI
 790

Fig. 6 SHEET 2

10 20 30 40
 MVYTLSGVRFPTVPSVYKSNNGFSSNGDRRNANVSVFLKKH--SLSRKILA
 MVYT: SG: RFP.: PS.: KS : . DRR.: : S FLK.: : S: SR. L
 MVYTI SGIRFPVLP SLHKS---TLRCDRRASSHSFFLKNSSSF SRTSLY
 10 20 30 40
 50 60 70 80 90
 EKSSYNSEFRPSTVAASGKVLVPGTQSDSSSSSTDOFEFTETSPENSPAS
 . K S : SE : : ST: A. S: KVL: P. . Q D: S S : DQ: E : E: . . .
 AKFSRDSETKSSTIAESDKVLIPEDQ-DNSVSLADQLENPDITSEDAQNL
 50 60 70 80 90
 100 110 120 130 140
 TDVDSSTMEHASQIKTENDDVEPSSDLTGSVEELDFASSLQLOEGGKLEE
 . D: TM. : : : : : : : : : : : : : : : : S : : : : : : : :
 EDL---TMKDGKNKYNI D-ESTSSYREVGDEKGSVTSSSLVDVNTDTQ--A
 100 110 120 130 140
 150 160 170 180 190
 SKTLNTSEETI IDESDRIRERGI PPPGLGQKIYEIDPLL TNYROHLDYRY
 . KT S: . . . : . : I IPPPG GQKIYEIDPLL . . ROHLD: RY
 KKTSVHSDKKVKVDKPKI----IPPPGSGQKIYEIDPLLQAHROHLD FRY
 150 160 170 180
 200 210 220 230 240
 SOYKKLREAI DKYEGGLEAFSRGYEKMGTFRSATGITYREWALGAQSAAL
 : QYK: : RE. IDKYEGGL: AFSRGYEK. GFTRSATGITYREW: GA: SAAL
 GOYKRIREE IDKYEGGLDAFSRGYEKFGFTRSATGITYREWGP GAKSAAL
 190 200 210 220 230
 250 260 270 280 290
 IGDFNNWDANADIMTRNEFGVWEIFLPNNVDGSPAIPHGSRVKIRMDTPS
 : GDFNNW: : NAD: MT: : . FGVWEIFLPNN. DGSP: IPHGSRVKI: MDTPS
 VGDFNNWNPNADVMTKDAFGVWEIFLPNNADGSPPIPHGSRVKIHMDTPS
 240 250 260 270 280
 300 310 320 330 340
 GVKDSIPAWINYSLQLPDEIPYNGIHYDPPEEERYIFQHPRPKPKSLRI
 G: KDSIPAWI: : S: Q P: EIPYNGI. YDPPEEE: Y: F: HP: PK: P: S: RI
 GIKDSIPAWIKFSVQAPGEIPYNGIYYDPPEEEKYVFKHPQPKRPOSIRI
 290 300 310 320 330
 350 360 370 380 390
 YESHIGMSSPEPKINSYVNRDEVLPRIKKLGYNALQIMAIQEHSYYASF
 YESHIGMSSPEPKIN: Y. NFRD: VLPRIKKLGYN: QIMAIQEHSYYASF
 YESHIGMSSPEPKINTYANFRDDVLPRIKKLGYNVQIMAIQEHSYYASF
 340 350 360 370 380
 400 410 420 430 440
 GYHVTNFFAPSSRFGTPDDLKSLIDKAHELGI VVLM DIVHSHASNNTLDG
 GYHVTNFFAPSSRFGTP: DLKSLID: AHEL G: : VLM DIVHSH: SNNTLDG
 GYHVTNFFAPSSRFGTPEDLKSLIDRAHELGLLVLM DIVHSHSSNNTLDG
 390 400 410 420 430

Fig. 7 SHEET 1

23/75

↙450 ↙460 ↙470 ↙480 ↙490
 LNMFDCTDSCYFHSGARGYHWMWDSRLFNNGNWEVRLRYLLSNARWWLDAF
 LNMFD TD: YFH: G: RGYHWMWDSRLFNNG: WEVRLRYLLSNARWWLD.:
 LNMFDGTDGHYFHPGSRGYHWMWDSRLFNNGSWEVRLRYLLSNARWWLDEY
 ^440 ^450 ^460 ^470 ^480
 ↙500 ↙510 ↙520 ↙530 ↙540
 KFDGFRFDGVTSMYIHHGLSVGFTGNYEEYFGLATDVEDAVVYMLVNDL
 KFDGFRFDGVTSMY. HHGL V: FTGNY. EYFGLATDV: AVVY: MLVNDL
 KFDGFRFDGVTSMYTHHGLQVSFTGNYSEYFGLATDVEAVVYMLVNDL
 ^490 ^500 ^510 ^520 ^530
 ↙550 ↙560 ↙570 ↙580 ↙590
 IHGLFPDAITIGEDVSGMPTFCIPVQEGGVGFDYRLHMAIADKRIELLKK
 IHGLFP: A: : IGEDVSGMPTFC: P. Q: GG: GF: YRLHMA: ADK: IELLKK
 IHGLFPEAVSIGEDVSGMPTFCLPTODGGIGFNRYRLHMAVADKWIELLKK
 ^540 ^550 ^560 ^570 ^580
 ↙600 ↙610 ↙620 ↙630 ↙640
 RDEDWRVGDIVHTLTNRRWSEKCVSYAESHDQALVGDKTIAFWLMDKDMY
 : DEDWR: GDIVHTLTNRRW EKC V YAESHQALVGDKT: AFWLMDKDMY
 QDEDWRMGDIVHTLTNRRWLEKCVVYAESHQALVGDKTLAFWLMDKDMY
 ^590 ^600 ^610 ^620 ^630
 ↙650 ↙660 ↙670 ↙680 ↙690
 DFMALDRPSTSLIDRGIALHKMIRLVTMGLGGEGYLNFMGNEFGHPEWID
 DFMALDRPST: LIDRGIALHKMIRL: TMGLGGEGYLNFMGNEFGHPEWID
 DFMALDRPSTPLIDRGIALHKMIRLITMGLGGEGYLNFMGNEFGHPEWID
 ^640 ^650 ^660 ^670 ^680
 ↙700 ↙710 ↙720 ↙730 ↙740
 FPRAEQHLSGGSVIPGNQFSYDKCRRRFDLGDAEYLYRGLQEFDRPMQY
 FPR: EQHL: : G. : : PGN: SYDKCRRRFDLGDA: YLRY: G: QEFDR: MQ.
 FPRGEQHLPNGKIVPGNNNSYDKCRRRFDLGDAEYLYRHGMQEFDRAMQH
 ^690 ^700 ^710 ^720 ^730
 ↙750 ↙760 ↙770 ↙780 ↙790
 LEDKYEFTMTSEHQFISRKDEGDRMIVFEKGNLVFVFNFWHTKSYSDYRIA
 LE: . Y. FMTSEHQ: ISRK: EGDR: I: FE: : NLVFNFWHT: SYSDY: : :
 LEETYGFMTSEHQYISRKNEGDRVIFERDNLVFVFNFWHTNSYSDYKVG
 ^740 ^750 ^760 ^770 ^780
 ↙800 ↙810 ↙820 ↙830 ↙840
 CLKPGKYKVALDSDDPLFGGFRIDHNAEYFTFEGWYDDRPRSIMVYAPC
 CLKPGKYK: . LDSDD. LFGGF. R: : H. AEYFT EGWYDDRPRS: : VYAP.
 CLKPGKYKIVLDSDDTLFGGFNRLNHTAEYFTSEGWYDDRPRSFLVYAPS
 ^790 ^800 ^810 ^820 ^830
 ↙850 ↙860 ↙870
 KTAVVYALVDKEEEEEEEEEEEVAA
 : TAVVYAL. D E. E E. : . V. :
 RTAVVYALADGVESEPIELSDGVES
 ^840 ^850 ^860

Fig. 7 SHEET 2

24 / 75

1 -----TTG-AT-----
 1 -----TTGA-----
 1 -----GA-----
 45 **AAAAACCTCCTCCACTCAGTCTTGGAATCTCTCTCTCT**

 72 TTTCTCTTAATTCCAACCA**GGGA**ATGAATAAAAGGAT-A
 73 TTTCTCTTAATTCCAACCAAGG-AATGAATAAAAGGAT-A
 71 TTTCTCTTAATTCCAACCAAGG-AATGAATAAAA**AGAT**-A
 165 TTTCTCTTAATTCCAACCAAGG-AATGAAT**IAAAAAGATIA**

 191 TGTACAAATCTAATGGATT**CAGCAGTAATGGTGATCGGAG**
 191 TGTACAAATCTAATGGATT**CAGCAGTAATGGTGATCGGAG**
 189 TGTACAAATCTAATGGATT**CAGCAGTAATGGTGATCGGAG**
 274 TGTACAAATCTAATGGATT**CAGCAGTAATGGTGATCGGAG**

 311 AATCCGACCTTCTACAGTTGCAGCATCGGGGAAAGTCCT
 311 AATCCGACCTTCTACAGTTGCAGCATCGGGGAAAGTCCT
 309 AAT**CCGACCTTCTACA**ATTGCAGCATCGGGGAAAGTCCT
 394 AAT**CCGACCTTCTACA**ATTGCAGCATCGGGGAAAGTCCT

 431 CAGCATCAACTGATGTAGATAGTTCAACAATGGAACACGC
 431 CAGCATCAACTGATGTAGATAGTTCAACAATGGAACACGC
 429 CAGCATCAACTGATGTAGATAGTTCAACAATGGAACACGC
 514 CAGCATCAACTGATGT**G**ATAGTTCAACAATGGAACACGC

 551 CATCACTACAAC**TACAAGAAGGTGGTAAACTGGAGGAGTC**
 551 CATCACTACAAC**TACAAGAAGGTGGTAAACTGGAGGAGTC**
 549 CATCACTACAAC**TACAAGAAGGTGGTAAACTGGAGGAGTC**
 634 CATCACTACAAC**TACAAGAAGGTGGTAAACTGGAGGAGTC**

 671 TTGGTCAGAAGATTTATGAAATAGACCCCTTTTGACAAA
 671 TTGGTCAGAAGATTTATGAAATAGACCCCTTTTGACAAA
 669 TTGGTCAGAAGATTTATGAAATAGACCCCTTTTGACAAA
 754 TTGGTCAGAAGATTTATGAAATAGACCCCTTTTGACAAA

 791 AAGC**TTTTCTCGTGGTTATGAAAAAATGGGTTTCACTCG**
 791 AAGC**TTTTCTCGTGGTTATGAAAAAATGGGTTTCACTCG**
 789 AAGCTTTTTCTCGTGGTTATGAA**A**GAATGGGTTTCACTCG
 874 AAGCTTTTTCTCGTGGTTATGAAAAAATGGGTTTCACTCG

Fig.8
Sheet 2

25/75

-----GGGCCTTGAAGTCAAGCAATTTGACACTCAGTTAGTTAC
-----TGGGGCCTTGAAGTCAAGCAATTTGACACTCAGTTAGTTAC
-----TGGGGCCTTGAAGTCAAGCAATTTGACACTCAGTTAGTTAC
TCACGCTTCTCTTGGGGCCTTGAAGTCAAGCAATTTGACACTCAGTTAGTTAC

GATTTGTAAAAACCTAAGGAGAGAAGAAGAAAGATGGTGTATA**A**ACTCTCT
GATTTGTAAAAACCTAAGGAGAGAAGAAGAAAGATGGTGTATACACTCTCT
GATTTGTAAAAACCTAAGGAGAGAAGAAGAAAGATGGTGTATACACTCTCT
GATTTG-----AAGGAGAGAAGAAGAAAGATGGTGTATACACTCTCT

GAATGCTAATGTTTCTGTATTCTTGAAAAAGCACTCTCTTTCACGGAAGATC
GAATGCTAATGTTTCTGTATTCTTGAAAAAGCACTCTCTTTCACGGAAGATC
GAATGCTAAT**A**TTTCTGTATTCTTGAAAA**A**CACTCTCTTTCACGGAAGATC
GAATGCTAATGTTTCTGTATTCTTGAAAAAGCACTCTCTTTCACGGAAGATC

TGTGCCTGGAA**C**CCAGAGTGATAGCTCCTCATCCTCAACAGACCAATTTGAG
TGTGCCTGGAA**C**CCAGAGTGATAGCTCCTCATCCTCAACAGACCAATTTGAG
TGTGCCTGGAA**T**CCAGAGTGATAGCTCCTCATCCTCAACAG**A**CAATTTGAG
TGT**A**CCTGGAATCCAGAGTGATAGCTCCTCATCCTCAACAGACCAATTTGAG

Fig. 8
Sheet
3

TAGCCAGATTAAAACTGAGAACGATGACGTTGAGCCGTCAAGTGATCTTACA
TAGCCAGATTAAAACTGAGAACGATGACGTTGAGCCGTCAAGTGATCTTACA
TAGCCAGATTAAAACTGAGAACGATGACGTTGAGCCGTCAAGTGATCTTACA
TAGCCAGATTAAAACTGAGAACGATGACGTTGAGCCGTCAAGTGATCTTACA

TAAAACATTAAATACTTCTGAAGAGACAATTATTGATGAATCTGATAGGATC
TAAAACATTAAATACTTCTGAAGAGACAATTATTGATGAATCTGATAGGATC
TAAAACATTAAATACTTCTGAAGAGACAATTATTGATGAATCTGATAGGATC
TAAAACATTAAATACTTCTGAAGAGACAATTATTGATGAATCTGATAGGATC

CTATCGTCAACACCTTGATTACAGGTATTCACAGTACAAGAACTGAGGGAG
CTATCGTCAACACCTTGATTACAGGTATTCACAGTACAAGAACTGAGGGAG
CTATCGTCAACACCTTGATTACAGGTATTCACAGTACAAGAACTGAGGGAG
CTATCGTCAACACCTTGATTACAGGTATTCACAGTACAAGAAA**A**TGAGGGAG

TAGTGCTACAGGTATCACTTACCGTGAGTGGGCTCCTGGTGCCAGTCAGCT
TAGTGCTACAGGTATCACTTACCGTGAGTGGGCT**C**TGGTGCCAGTCAGCT
TAGTGCTACAGGTATCACTTACCGTGAGTGGGCTCCTGGTGCCAGTCAGCT
TAGTGCTACAGGTATCACTTACCGTGAGTGGGCTCCTGGTGCCAGTCAGCT

Fig. 8 SHEET 2

26/75

ACTCCTATCACTTATCAGATCTCTATTT 11con.seq
ACTCCTATCACTTATCAGATCTCTATTT 19con.seq
ACTGCTATCACTTATCAGATCTCTATTT 10con.seq
ACTCCTATCACTGATCAGATCTCTATTT psbe2con.seq

GGAGTTCGTTTTCTACTGTTCCATCAG 11con.seq
GGAGTTCGTTTTCTACTGTTCCATCAG 19con.seq
GGAGTTCGTTTTCTACTGTTCCATCAG 10con.seq
GGAGTTCGTTTTCTACTGTTCCATCAG psbe2con.seq

TTGGCTGAAAAGTCTTCTTACAATTCCG 11con.seq
TTGGCTGAAAAGTCTTCTTACAATTCCG 19con.seq
TTGGCTGAAAAGTCTTCTTACAATTCCG 10con.seq
TTGGCTGAAAAGTCTTCTTACGATTCCG psbe2con.seq

TTCACTGAGACATCTCCAGAAAATTCCC 11con.seq
TTCAGTCTGAGACATCTCCAGAAAATTCCC 19con.seq
TTCGCTGAGACATCTCCAGAAAATTCCC 10con.seq
TTCAGTCTGAGACAGCTCCAGAAAATTCCC psbe2con.seq

GGAAGTGTTGAAGAGCTGGATTTTGCTT 11con.seq
GGAAGTGTTGAAGAGCTGGATTTTGCTT 19con.seq
GGAAGTGTTGAAGAGCTGGATTTTGCTT 10con.seq
GGAAGTGTTGAAGAGCTGGATTTTGCTT psbe2con.seq

AGAGAGAGGGGCATCCCTCCACCTGGAC 11con.seq
AGAGAGAGGGGCATCCCTCCACCTGGAC 19con.seq
AGAGAGAGGGGCATCCCTCCACCTGGAC 10con.seq
AGAGAGAGGGGCATCCCTCCACCTGGAC psbe2con.seq

GCAATTGACAAGTATGAGGGTGGTTTGG 11con.seq
GCAATTGACAAGTATGAGGGTGGTTTGG 19con.seq
GCAATTGACAAGTATGAGGGTGGTTTGG 10con.seq
GCAATTGACAAGTATGAGGGTGGTTTGG psbe2con.seq

GCCCTCATTGGAGATTTCACAATTGGG 11con.seq
GCCCTCATTGGAGATTTCACAATTGGG 19con.seq
GCCCTCATTGGGATTTCACAATTGGG 10con.seq
GCCCTCATTGGGATTTCACAATTGGG psbe2con.seq

Fig. 8
SHEET 3

910 ACGCAAATGCTGACATTATGACTCGGAATGAATTTGGTGTC
 911 ACGCAAATGCTGACATTATGACTCGGAATGAATTTGGTGTC
 909 ACGCAAATGCTGACATTATGACTCGGAATGAATTTGGTGTC
 994 ACGCAAATGCTGACATTATGACTCGGAATGAATTTGGTGTC

 1030 CTCCATCAGGTGTTAAGGATTCCATTCTGCTTGGATCAAC
 1031 CTCCATCAGGTGTTAAGGATTCCATTCTGCTTGGATCAAC
 1029 CTCCATCAGGTGTTAAGGATTCCATTCTGCTTGGATCAAC
 1114 CTCCATCAGGTGTTAAGGATTCCATTCTGCTTGGATCAAC

 1150 AACACCCACGGCCAAAGAAACCAAAGTCGCTGAGAATATAT
 1151 AACACCCACGGCCAAAGAAACCAAAGTCGCTGAGAATATAT
 1149 AACACCCACGGCCAAAGAAACCAAAGTCGCTGAGAATATAT
 1234 AACACCCACGGCCAAAGAAACCAAAGTCGCTGAGAATATAT

 1270 TAAAAAA-GCTTGGGTACAATGCGCTGCAATTATGGCTAT
 1271 TAAAAAA-GCTTGGGTACAATGCGCTGCAATTATGGCTAT
 1269 TAAAAAAAGCTTGGGTACAATGCGGTGCAATTATGGCTAT
 1354 TAAAAAAC-CTTGGGTACAATGCGGTGCAATTATGGCTAT

 1389 GACGACCTTAAGTCTTGATTGATAAAGCTCATGAGCTAGG
 1390 GACGACCTTAAGTCTTTGATTGATAAAGCTCATGAGCTAGG
 1389 GACGACCTTAAGTCTTTGATTGATAAAGCTCATGAGCTAGG
 1473 GACGACCTTAAGTCTTTGATTGATAAAGCTCATGAGCTAGG

 1509 GATAGTTGTTACTTTCACTCTGGAGCTCGTGGTTATCATTG
 1510 GATAGTTGTTACTTTCACTCTGGAGCTCGTGGTTATCATTG
 1509 GATAGTTGTTACTTTCACTCTGGAGCTCGTGGTTATCATTG
 1593 GATAGTTGTTACTTTCACTCTGGAGCTCGTGGTTATCATTG

 1628 GATGAGTTCAAATTTGATGGATTAGATTGATGGTGTGAC
 1630 GATGAGTTCAAATTTGATGGATTAGATTGATGGTGTGAC
 1629 GATGAGTTCAAATTTGATGGATTAGATTGATGGTGTGAC
 1713 GATGAGTTCAAATTTGATGGATTAGATTGATGGTGTGAC

 1748 GTGGATGCTGTTGTGTATCTGATGCTGGTCAACGATCTTAT
 1750 GTGGATGCTGTTGTGTATCTGATGCTGGTCAACGATCTTAT
 1749 GTGGATGCTGTTGTGTATCTGATGCTGGTCAACGATCTTAT
 1833 GTGGATGCTGTTGTGTATCTGATGCTGGTCAACGATCTTAT

Fig. 8
Sheet 5

Fig. 8
SHEET 4

TGGGAGATTTTTCTGCCAAATAATGTGGATGGTTCTCCTGCAATTC
 TGGGAGATTTTTCTGCCAAATAATGTGGATGGTTCTCCTGCAATTC
 TGGGAGATTTTTCTGCCAAATAATGTGGATGGTTCTCCTGCAATTC
 TGGGAGATTTTTCTGCCAAATAATGTGGATGGTTCTCCTGCAATTC

TACTCTTTACAGCTTCCTGATGAAATTCATATAATGGAATATATT
 TACTCTTTACAGCTTCCTGATGAAATTCATATAATGGAATATATT
 TACTCTTTACAGCTTCCTGATGAAATTCATATAATGGAATATATT
 TACTCTTTACAGCTTCCTGATGAAATTCATATAATGGAATATATT

GAATCTCATATTGGAATGAGTAGTCCGGAGCCTAAAATTAACATCAT
 GAATCTCATATTGGAATGAGTAGTCCGGAGCCTAAAATTAACATCAT
 GAATCTCATATTGGAATGAGTAGTCCGGAGCCTAAAATTAACATCAT
 GAATCTCATATTGGAATGAGTAGTCCGGAGCCTAAAATTAACATCAT

TCAAGAGCATTCTTATTATGCTAGTTTTGGTTATCATGTCACAAAT
 TCAAGAGCATTCTTATTATGCTAGTTTTGGTTATCATGTCACAAAT
 TCAAGAGCATTCTTATTATGCTAGTTTTGGTTATCATGTCACAAAT
 TCAAGAGCATTCTTATTATGCTAGTTTTGGTTATCATGTCACAAAT

AATTGTTGTTCTCATGGACATTGTTACAGCCATGCATCAAATAAT
 AATTGTTGTTCTCATGGACATTGTTACAGCCATGCATCAAATAAT
 AATTGTTGTTCTCATGGACATTGTTACAGCCATGCATCAAATAAT
 AATTGTTGTTCTCATGGACATTGTTACAGCCATGCATCAAATAAT

GATGTGGGATTCCGCCTCTTTAACTATGGAACTGGGAGGTACTT
 GATGTGGGATTCCGCCTCTTTAACTATGGAACTGGGAGGTACTT
 GATGTGGGATTCCGCCTCTTTAACTATGGAACTGGGAGGTACTT
 GATGTGGGATTCCGCCTCTTTAACTATGGAACTGGGAGGTACTT

ATCAATGATGTATACTCACCACGGATTATCGGTGGGATTCACTGGG
 ATCAATGATGTATACTCACCACGGATTATCGGTGGGATTCACTGGG
 ATCAATGATGTATACTCACCACGGATTATCGGTGGGATTCACTGGG
 ATCAATGATGTATACTCACCACGGATTATCGGTGGGATTCACTGGG

TCATGGGCTTTTCCCAGATGCAATTACCATTGGTGAAGATGTTAGC
 TCATGGGCTTTTCCCAGATGCAATTACCATTGGTGAAGATGTTAGC
 TCATGGGCTTTTCCCAGATGCAATTACCATTGGTGAAGATGTTAGC
 TCATGGGCTTTTCCCAGATGCAATTACCATTGGTGAAGATGTTAGC

Fig. 8
 Sheet 6

Fig. 8
 SHEET 5

29/75

CTCATGGGTCCAGAGTGAAGATACGTATGGACA 11con.seq
CTCATGGGTCCAGAGTGAAGATACGTATGGACA 19con.seq
CTCATGGGTCCAGAGTGAAGATACGTATGGACA 10con.seq
CTCATGGGTCCAGAGTGAAGATACGATGGACA psbe2con.seq

ATGATCCACCCGAAGAGGAGAGGTATATCTTCC 11con.seq
ATGATCCACCCGAAGAGGAGAGGTATATCTTCC 19con.seq
ATGATCCACCCGAAGAGGAGAGGTATATCTTCC 10con.seq
ATGATCCACCCGAAGAGGAGAGGTATCTTCC psbe2con.seq

ACGTGAATTTTAGAGATGAAGTTCTTCCTCGCA 11con.seq
ACGTGAATTTTAGAGATGAAGTTCTTCCTCGCA 19con.seq
ACGTGAATTTTAGAGATGAAGTTCTTCCTCGCA 10con.seq
ACGTGAATTTTAGAGATGAAGTTCTTCCTCGCA psbe2con.seq

TTTTTTGCACCAAGCAGCCGTTTTGGAACGCCC 11con.seq
TTTTTTGCACCAAGCAGCCGTTTTGGAACGCCC 19con.seq
TTTTTTGCACCAAGCAGCCGTTTTGGAACGCCC 10con.seq
TTTTTTGCACCAAGCAGCCGTTTTGGAACGCCC psbe2con.seq

ACTTTAGATGGACTGAACATGTTTGACGGCACC 11con.seq
ACTTTAGATGGACTGAACATGTTTGACGGCACC 19con.seq
ACTTTAGATGGACTGAACATGTTTGACGGCACA 10con.seq
ACTTTAGATGGACTGAACATGTTTGACGGCACA psbe2con.seq

AGGTATCTTCTCTCAAATGCGAGATGGTGGTTG 11con.seq
AGGTATCTTCTCTCAAATGCGAGATGGTGGTTG 19con.seq
AGGTATCTTCTCTCAAATGCGAGATGGTGGTTG 10con.seq
AGGTATCTTCTCTCAAATGCGAGATGGTGGTTG psbe2con.seq

AACTACGAGGAATACTTTGGACTCGCAACTGAT 11con.seq
AACTACGAGGAATACTTTGGACTCGCAACTGAT 19con.seq
AACTACGAGGAATACTTTGGACTCGCAACTGAT 10con.seq
AACTACGAGGAATACTTTGGACTCGCAACTGAT psbe2con.seq

GGAATGCCGACATTTTGTATTCCCGTTCAAGAT 11con.seq
GGAATGCCGACATTTTGTATTCCCGTTCAAGAT 19con.seq
GGAATGCCGACATTTTGTATTCCCGTTCAAGAT 10con.seq
GGAATGCCGACATTTTGTATTCCCGTTCAAGAT psbe2con.seq

Fig. 8
SHEET 6

1868 GGGGGTGTGGCTTTGACTATCGGCTGCATATGGCAATTGC
1870 GGGGGTGTGGCTTTGACTATCGGCTGCATATGGCAATTGC
1869 GGGGGTGTGGCTTTGACTATCGGCTGCATATGGCAATTGC
1953 GGGGGTGTGGCTTTGACTATCGGCTGCATATGGCAATTGC

1988 AGATGGTCGGAAAAGTGTGTTTCATACGCTGAAAGTCATGA
1990 AGATGGTCGGAAAAGTGTGTTTCATACGCTGAAAGTCATGA
1989 AGATGGTCGGAAAAGTGTGTTTCATACGCTGAAAGTCATGA
2073 AGATGGTCGGAAAAGTGTGTTTCATACGCTGAAAGTCATGA

2108 CCGCAACATCATTAATAGATCGTGGGATAGCATTGCACAA
2110 CCGTCAACATCATTAATAGATCGTGGGATAGCATTGCACAA
2109 CCGTCAACATCATTAATAGATCGTGGGATAGCATTACACAA
2193 CCGTCAACATCATTAATAGATCGTGGGATAGCATTGCACAA

2228 TGGATTGATTTCCCTAGGGCTGAACACACCTTCTCTGATGG
2230 TGGATTGATTTCCCTAGGGCTGAACAACACCTCTCTGATGG
2229 TGGATTGATTTCCCTAGGGCTGAACAACACCTCTCTGATGG
2313 TGGATTGATTTCCCTAGGGCTGAACAACACCTCTCTGATGG

2348 TACCAATGGGTTACAAGAATTTGACTGGGCTATGCAGTATCT
2350 TACCGTGGGTTGCAAGAATTTGACCGGCTATGCAGTATCT
2349 TACCGTGGGTTGCAAGAATTTGACCGGGCTATGCAGTATCT
2433 TACCGTGGGTTGCAAGAATTTGACCGGGCTATGCAGTATCT

2468 GAAA**G**AGGAAACCTAGTTTT**C**GTCTTTAATTTTCACTGGAC
2470 GAAAAAGGAAACCTAGTTTTGTCTTTAATTTTCACTGGAC
2469 GAAAAAGGAAACCTAGTTTTGTCTTTAATTTTCACTGGAC
2553 GAAAAAGGAAACCTAGTTTTGTCTTTAATTTTCACTGGAC

2588 TTTGGTGGCTTCGGGAGAATTGATCATAATGCCGAATATTT
2590 TTTGGTGGCTTCGGGAGAATTGATCATAATGCCGAATATTT
2589 TTTGGTGGCTTCGGGAGAATTGATCATAATGCCGAATATTT
2673 TTTGGTGGCTTCGGGAGAATTGATCATAATGCCGAATGTTT

2708 CTAGTAGACAAA**CT**AGAAG-----
2710 CTAGTAGACAAAGAAGAAGAAGAAGAAG**AAGAAGA**
2709 CTAGTAGACAAAGAAGAAGAAGAAGAAGAAG-----
2793 CTAGTAGACAAAGAAGAAGAAGAAGAAG**---**-----

Fig.8
Sheet 8

Fig. 8
SHEET 7

31/75

TGATAAATGGATTGAGTTGCTCAAGAAACGGGATGAGGATTGGAGA
TGATAAA[GGATTGAGTTGCTCAAGAAACGGGATGAGGATTGGAGA
TGATAAATGGATTGAGTTGCTCAAGAAACGGGATGAGGATTGGAGA
TGATAAATGGATTGAGTTGCTCAAGAAACGGGATGAGGATTGGAGA

TCAAGCTCTAGTCGGTGATAAACTATAGCATTCTGGCTGATGGAC
TCAAGCTCTAGTCGGTGATAAACTATAGCATTCTGGCTGATGGAC
TCAAGCTCTAGTCGGTGATAAACTATAGCATTCTGGCTGATGGAC
TCAAGCTCTAGTCGGTGATAAACTATAGCATTCTGGCTGATGGAC

GATGATTAGGCTTGTAACCTATGGGATTAGGAGGAGAAGGGTACCTA
GATGATTAGGCTTGTAACCTATGGGATTAGGAGGAGAAGGGTACCTA
GATGATTAGGCTTGTAACCTATGGGATTAGGAGGAGAAGGGTACCTA
GATGATTAGGCTTGTAACCTATGGGATTAGGAGGAGAAGGGTACCTA

CTCAGTAATTC[CCCGGAAACCAATTCAGTTATGATAAATGCAGACGG
CTCAGTAAT[CCCGGAAACCAATTCAGTTATGATAAATGCAGACGG
CTCAGTAATTC[CCCGGAAACCAATTCAGTTATGATAAATGCAGACGG
CTCAGTAATTC[CCCGGAAACCAATTCAGTTATGATAAATGCAGACGG

TGAAGATAAATATGAGTTTATGACTTCAGAACACCAGTTCATATCA
TGAAGATAAATATGAGTTTATGACTTCAGAACACCAGTTCATATCA
TGAAGATAAATATGAGTTTATGACTTCAGAACACCAGTTCATATCA
TGAAGATAAATATGAGTTTATGACTTCAGAACACCAGTTCATATCA

AAA[AGCTATTCAGACTATCGCATAGGCTGCCTGAAGCCTGGAAAA
AAAAAGCTATTCAGACTATCGCATAG[CTGCCTGAAGCCTGGAAAA
AAAA[GGCTATTCAGACTATCGCATAGGCTGCCTGAAGCCTGGAAAA
AAAAAGCTATTCAGACTATCGCATAGGCT[CTGAAGCCTGGAAAA

CACCT[CTGAAGGAT[GTATGATGATCGTCCT[GTTCATTATGGTG
CACCTTTGAAGGATGGTATGATGATCGTCCTCGTTCAATTATGGTG
CACCTTTGAAGGATGGTATGATGATCGTCCTCGTTCAATTATGGTG
CACCTTTGAAGGATGGTATGATGATCGTCCTCGTTCAATTATGGTG

-----TAGCAGTAGTAGAAGAA[CCCATTTG-----AAGAATGAACG
AGAAGTAGCAG[CAGTAGAAGAAGTAGTAGTAGAAGAAGAATGAACG
-----TAGCAGTAGTAGAAGAAGTAGTAGTAGAAGAAGAATGAACG
-----TAGCAGTAGTAGAAGAAGTAGTAGTAGAAGAAGAATGAACG

Fig.8
Sheet 9

Fig. 8
SHEET 8

32/75

GTGGGTGATATTGTTCATACACTGACAAATAGA 11con.seq
GTGGGTGATATTGTTCATACACTGACAAATAGA 19con.seq
GTGGGTGATATTGTTCATACACTGACAAATAGA 10con.seq
GTGGGTGATATTGTTCATACACTGACAAATAGA psbe2con.seq

AAGGATATGTATGATTTTATGGCTCTGGATAGA 11con.seq
AAGGATATGTATGATTTTATGGCTCTGGATAGA 19con.seq
AAGGATATGTATGATTTTATGGCTCTGGATAGA 10con.seq
AAGGATATGTATGATTTTATGGCTTTGGATAGA psbe2con.seq

AATTTTCATGGGAAATGAATTCGGCCACCCTGAG 11con.seq
AATTTTCATGGGAAATGAATTCGGCCACCCTGAG 19con.seq
AATTTTCATGGGAAATGAATTCGGCCACCCTGAG 10con.seq
AATTTTCATGGGAAATGAATTCGGCCACCCTGAG psbe2con.seq

AGATTTGACCTGGGAGATGCAGAATATTTAAGA 11con.seq
AGATTTGACCTGGGAGATGCAGAATATTTAAGA 19con.seq
AGATTTGACCTGGGAGATGCAGAATATTTAAGA 10con.seq
AGATTTGACCTGGGAGATGCAGAATATTTAAGA psbe2con.seq

CGAAAGGATGAAGGAGATAGGATGATTGTATTT 11con.seq
CGAAAGGATGAAGGAGATAGGATGATTGTATTT 19con.seq
CGAAAGGATGAAGGAGATAGGATGATTGTATTT 10con.seq
CGAAAGGATGAAGGAGATAGGATGATTGTATTT psbe2con.seq

TACAAGGTTCTTGGACTCAGATGATCCACTT 11con.seq
TACAAGGTTGCCTTGGACTCAGATGATCCACTT 19con.seq
TACAAGGTTGCCTTGGACTCAGATGATCCACTT 10con.seq
TACAAGGTTGCCTTGGACTCAGATGATCCACTT psbe2con.seq

TATGCACCTAGTAGAACAGCAGTGGTCTATGCA 11con.seq
TATGCACCTTGTAAACAGCAGTGGTCTATGCA 19con.seq
TATGCACCTAGTAGAACAGCAGTGGTCTATGCA 10con.seq
TATGCACCTAGTAGAACAGCAGTGGTCTATGCA psbe2con.seq

AACTTGTGATCGCGTTGAAAGATTTGAACGTTA 11con.seq
AACTTGTGATCGCGTTGAAAGATTTGAACG--- 19con.seq
AACTTGTGATCGCGTTGAAAGATTTGAACG--- 10con.seq
AACTTGTGATCGCGTTGAAAGATTTGAACG--- psbe2con.seq

Fig. 8
SHEET 9

33/75

2795 CTTGGTCATCCACATAGAGCTTCTTGAC-----
2827 -----CTACATAGAGCTTCTTGACGTATCTGGCAATAT
2814 -----CCACATAGAGCTTCTTGACGTATCTGGCAATAT
2895 -----CTACATAGAGCTTCTTGACGTATCTGGCAATAT

2898 AGAGATGAAGTGCTGAACAAA--CATATGTAAAATCGATGAA
2937 AGAGATGAAGTGCTGAACAAA--CATATGTAAAATCGATGAA
2924 AGAGATGAAGTGCTGAACAAAACATATGTAAAATCGATGAA
3005 AGAGATGAAGTGCTGAACAAA--CATATGTAAAATCGATGAA

2975
3012
3003
3123 GCCCACTAGAAATCAATTATGTGAGACCTAAAAACAATAAC

Fig. 8
Sheet 11

Fig. 8 SHEET 10

34/75

---ATCAGTCTTGGCGGAATTTCATGTGACAA--CAAGGTTTGCACTT
TGCATCAGTCTTGGCGGAATTTTCATGTGACAC--AAGGTTTGCAATT
TGCATTTAGTCTTGGCGGAATTTTCATGTGACAA--CAGGTTTGCAATT
TGCATCAGTCTTGGCGGAATTTTCATGTGACAA--AAGGTTTGCAATT

TTTATGTCGAATGCTGGGACGATCGAATTCCTGCAGCC
TTTATGTCGAATGCTGGGACGATCGAATTCCTGCAG
TTTATGTCGAATGCTGGGACGATCGAATTCCTGCAGCC
TTTATGTCGAATGCTGGGACGGGCTTCAGCAGCTTTGCTTAGTGA

CATAAAATGGAAATAGTGCTGATCTAATGATGTTTAAANCCNNNNA

Fig. 8
Sheet 12

Fig. 8 SHEET 11

35/75

CTTTCCACTATTAGTAGT**CCAC**CGATATACGC 11con.seq
CTTTCCACTATTAGTAGTGCAACGATATACGC 19con.seq
CTTTCCACTATTAGTAGTGCAACGATATACGC 10con.seq
CTTTCCACTATTAGTAGTGCAACGATATACGC psbe2con.seq

11con.seq

19con.seq

10con.seq

GTTCTGTAAATTGTCATCTCTTTANATGTACA psbe2con.seq

11con.seq

19con.seq

10con.seq

AAAAAAAAAAAAAACTCGAG

psbe2con.seq

Fig. 8 SHEET 12

36/75

GGATGCTAATGTTTCTGTATTCTTGAAAAAGCACTCTCTTTTCACGG

CCTACGATTACAAAGACATAAGAACTTTTTTCGTGAGAGAAAGTGCC

A N V S V F L K K H S L S R

TTCTACAGTTGCAGCATCGGGGAAAGTCCTTGTGCCTGGAAYCCAG

AAGATGTCAACGTCGTAGCCCCTTTCAGGAACACGGACCTTRGGTC

S T V A A S G K V L V P G ? Q

GACATCTCCAGAAAATTCCCCAGCATCAACTGATGTAGATAGTTCA

CTGTAGAGGTCTTTTAAGGGGTCGTAGTTGACTACATCTATCAAGT

T S P E N S P A S T D V D S S

TGAGCCGTCAAGTGATCTTACAGGAAGTGTTGAAGAGCTGGATTTT

ACTCGGCAGTTCACTAGAATGTCCTTCACAACCTTCTCGACCTAAAA

E P S S D L T G S V E E L D F

TAAACATTAAATACTTCTGAAGAGACAATTATTGATGAATCTGAT

ATTTTGTAATTTATGAAGACTTCTCTGTTAATAACTACTTAGACTA

K T L N T S E E T I I D E S D

Hinc II

GATTTATGAAATAGACCCCCTTTTGACAACTATCGTCAACACCTT

CTAAATACTTTATCTGGGGGAAAACCTGTTTGATAGCAGTTGTGGAA

I Y E I D P L L T N Y R Q H L

Fig.9
Sheet
2

Fig.9 SHEET 1

37/75

Bgl II

AAGATCTTGGCTGAAAAGTCTTCTTACAATTCCGAATCCCGACC 90
TTCTAGAACCGACTTTTTCAGAAGAATGTTAAGGCTTAGGGCTGG
K I L A E K S S Y N S E S R P

AGTGATAGCTCCTCATCCTCAACAGACCAATTTGAGTTCACTGA 180
TCACTATCGAGGAGTAGGAGTTGTCTGGTTAAACTCAAGTGACT
S D S S S S S T D Q F E F T E

ACAATGGAACACGCTAGCCAGATTAAACTGAGAACGATGACGT 270
TGTTACCTTGTGCGATCGGTCTAATTTTGA CTCTTGCTACTGCA
T M E H A S Q I K T E N D D V

GCTTCATCACTACA ACTACAAGAAGGTGGTAAACTGGAGGAGTC 360
CGAAGTAGTGATGTTGATGTTCTTCCACCATTGACCTCCTCAG
A S S L Q L Q E G G K L E E S

AGGATCAGAGAGAGGGGCATCCCTCCACCTGGACTTGGTCAGAA 450
TCCTAGTCTCTCTCCCCGTAGGGAGGTGGACCTGAACCAGTCTT
R I R E R G I P P P G L G Q K

GATTACAGGTATTACAGTACAAGAACTGAGGGAGGCAATTGA 540
CTAATGTCCATAAGTGTCATGTTCTTTGACTCCCTCCGTAACT
D Y R Y S Q Y K K L R E A I D

Fig. 9 SHEET 2

38/75

Hind III

CAAGTATGAGGGTGGTTTGGGAAGCTTTTTCTCGTGGTTATGAAAAA
GTTCACTACTCCCACCAAACCTTCGAAAAAGAGCACCAATACTTTTT
K Y E G G L E A F S R G Y E K

Pvu II

GGCTCCTGGTGCCAGTCAGCTGCCCTCATTGGAGATTTCAACAAT
CCGAGGACCACGGGTCAGTCGACGGGAGTAACCTCTAAAGTTGTTA
A P G A Q S A A L I G D F N N

CTGGGAGATTTTTCTGCCAAATAATGTGGATGGTTCTCCTGCAATT
GACCCTCTAAAAAGACGGTTTATTACACCTACCAAGAGGACGTAA
W E I F L P N N V D G S P A I

TGTTAAGGATTCCATTCTTGCTTGGATCAACTACTCTTTACAGCTT
ACAATTCCTAAGGTAAGGACGAACCTAGTTGATGAGAAATGTCGAA
V K D S I P A W I N Y S L Q L

AGAGGAGAGGTATRTCTTCCAACACCCACGGCCAAAGAAACCAAAG
TCTCCTCTCCATAYAGAAGGTTGTGGGTGCCGTTTCTTTGGTTTC
E E R Y ? F Q H P R P K K P K

Fig.9
Sheet
4

Fig.9 SHEET 3

39/75

ATGGGTTTCACTCGTAGTGCTACAGGTATCACTTACCGTGAGTG
TACCCAAAGTGAGCATCACGATGTCCATAGTGAATGGCACTCAC 630
M G F T R S A T G I T Y R E W

TGGGACGCAAATGCTGACATTATGACTCGGAATGAATTTGGTGT
ACCCTGCGTTTACGACTGTAATACTGAGCCTTACTTAAACCACA 720
W D A N A D I M T R N E F G V

CCTCATGGGTCCAGAGTGAAGATACGYATGGACACTCCATCAGG
GGAGTACCCAGGTCTCACTTCTATGCRTACCTGTGAGGTAGTCC 810
P H G S R V K I R M D T P S G

CCTGATGAAATTCCATATAATGGAATATATTATGATCCACCCGA
GGACTACTTTAAGGTATATTACCTTATATAATACTAGGTGGGCT 900
P D E I P Y N G I Y Y D P P E

TCGCTGAGAATATATGAATCTCATATTGGAATGAGTAGTCCGGA
AGCGACTCTTATATACTTAGAGTATAACCTTACTCATCAGGCCT 990
S L R I Y E S H I G M S S P E

Fig. 9 SHEET 4

40/75

Xmn I

GCCTAAAATTAACATCATACGTGAATTTTAGAGATGAAGTTCTTCCT

CGGATTTTAATTGAGTATGCACTTAAAATCTCTACTTCAAGAAGGA

P K I N S Y V N F R D E V L P

TCAAGAGCATTCTTATTATGCTAGTTTTGGTTATCATGTCACAAAT

AGTTCTCGTAAGAATAATACGATCAAAACCAATAGTACAGTGTTTA

Q E H S Y Y A S F G Y H V T N

GTCTTTGATTGATAAAGCTCATGAGCTAGGAATTGTTGTTCTCATG

CAGAACTAACTATTTTCGAGTACTCGATCCTTAACAACAAGAGTAC

S L I D K A H E L G I V V L M

GAACATGTTTGACGGCACAGATAGTTGTTACTTTCACTCTGGAGCT

CTTGTAACAACTGCCGTGTCTATCAACAATGAAAGTGAGACCTCGA

N M F D G T D S C Y F H S G A

AAACTGGGAGGTACTTAGGTATCTTCTCTCAAATGCGAGATGGTGG

TTTGACCCTCCATGAATCCATAGAAGAGAGTTTACGCTCTACCACC

N W E V L R Y L L S N A R W W

ATCAATGATGTATACTCACCACGGATTATCGGTGGGATTCACTGGG

TAGTTACTACATATGAGTGGTGCCTAATAGCCACCCTAAGTGACCC

S M M Y T H H G L S V G F T G

Fig.9
Sheet
6

Fig.9 SHEET 5

41/75

CGCATAAAAAASCTTGGGTACAATGCGGTGCAAATTATGGCTAT
GCGTATTTTTTSGAACCCATGTTACGCCACGTTTAATACCGATA
R I K ? L G Y N A V Q I M A I

1080

TTTTTTGCACCAAGCAGCCGTTTTGGAACGCCCGACGACCTTAA
AAAAAACGTGGTTCGTGCGCAAAACCTTGCGGGCTGCTGGAATT
F F A P S S R F G T P D D L K

1170

GACATTGTTACAGCCATGCATCAAATAATACTTTAGATGGACT
CTGTAACAAGTGTCGGTACGTAGTTTATTATGAAATCTACCTGA
D I V H S H A S N N T L D G L

1260

Sac I

CGTGGTTATCATTGGATGTGGGATTCCCGCCTCTTTAACTATGG
GCACCAATAGTAACCTACACCCTAAGGGCGGAGAAATTGATACC
R G Y H W M W D S R L F N Y G

1350

TTGGATGAGTTCAAATTTGATGGATTTAGATTTGATGGTGTGAC
AACCTACTCAAGTTTAACTACCTAAATCTAACTACCACTG
L D E F K F D G F R F D G V T

1440

AACTACGAGGAATACTTTGGACTCGCAACTGATGTGGATGCTGT
TTGATGCTCCTTATGAAACCTGAGCGTTGACTACACCTACGACA
N Y E E Y F G L A T D V D A V

1530

Fig. 9 SHEET 6

42/75

Hinc II

TGTGTATCTGATGCTGGTCAACGATCTTATTCACGGGCTTTTCCCA
ACACATAGACTACGACCAGTTGCTAGAATAAGTGCCCGAAAAGGGT
V Y L M L V N D L I H G L F P

TTGTATTCCCGTTCAAGATGGGGGTGTTGGCTTTGACTATCGGCTG
AACATAAGGGCAAGTTCTACCCCCACAACCGAAACTGATAGCCGAC
C I P V Q D G G V G F D Y R L

GGATGAGGATTGGAGAGTGGGTGATATTGTTTCATACACTGACAAAT
CCTACTCCTAACCTCTCACCCTACTATAACAAGTATGTGACTGTTTA
D E D W R V G D I V H T L T N

TCAAGCTCTAGTCGGTGATAAAACTATAGCATYCTGGCTGATGGAC
AGTTTCGAGATCAGCCACTATTTTGATATCGTARGACCGACTACCTG
Q A L V G D K T I A ? W L M D

ATTAATAGATCGTGGGATAGCATTGCACAAGATGATTAGGCTTGTA
TAATTATCTAGCACCCCTATCGTAACGTGTTCTACTAATCCGAACAT
L I D R G I A L H K M I R L V

Fig.9
Sheet
8

Fig.9 SHEET 7

43/75

GATGCAATTACCATTGGTGAAGATGTTAGCGGAATGCCGACATT 1620
CTACGTTAATGGTAACCACTTCTACAATCGCCTTACGGCTGTAA
D A I T I G E D V S G M P T F

Nde I

CATATGGCAATTGCTGATAAATGGATTGAGTTGCTCAAGAAACG 1710
GTATACCGTTAACGACTATTTACCTAACTCAACGAGTTCTTTGC
H M A I A D K W I E L L K K R

AGAAGATGGTCGGAAAAGTGTGTTTCATMCGCTGAAAGTCATGA 1800
TCTTCTACCAGCCTTTTTCACACAAAGTAKGCGACTTTCAGTACT
R R W S E K C V S ? A E S H D

Hinc II

AAGGATATGTATGATTTTATGGCTCTGGATAGACCGTCAACATC 1890
TTCCTATACATACTAAAATACCGAGACCTATCTGGCAGTTGTAG
K D M Y D F M A L D R P S T S

Asp 718

Kpn I

ACTATGGGATTAGGAGGAGAAGGGTACCTAAATTTTCATGGGAAA 1980
TGATACCCTAATCCTCCTCTTCCCATGGATTAAAGTACCCTTT
T M G L G G E G Y L N F M G N

Fig. 9 SHEET 8

44/75

EcoR I

TGAATTCGGCCACCCTGAGTGGATTGATTTCCCTAGGGCTGARCAA
 ACTTAAGCCGGTGGGACTCACCTAACTAAAGGGATCCCGACTYGT
 E F G H P E W I D F P R A E Q

Ssp I

TGATAAATGCAGACGGAGATTTGACCTGGGAGATGCAGAATATTTA
 ACTATTTACGTCTGCCTCTAACTGGACCCTCTACGTCTTATAAAT
 D K C R R R F D L G D A E Y L

TGAAGATAAATATGAGTTTATGACTTCAGAACACCAGTTCATATCA
 ACTTCTATTTATACTCAAATACTGAAGTCTTGTGGTCAAGTATAGT
 E D K Y E F M T S E H Q F I S

CCTAGTTTTTGTCTTTAATTTTCACTGGACAAATAGCTATTCAGAC
 GGATCAAAAACAGAAATTAAGGTGACCTGTTTATCGATAAGTCTG
 L V F V F N F H W T N S Y S D

GGACTCAGATGATCCACTTTTTGGTGGCTTCGGGAGAATTGATCAT
 CCTGAGTCTACTAGGTGAAAAACCACCGAAGCCCTCTTAAGTAGTA
 D S D D P L F G G F G R I D H

YCGYYCAATTATGGTGTATGCACCTAGTAGAACAGCAGTGGTCTAT
 RGCRRGTTAATACCACATACGTGGATCATCTTGTCGTCACCAGATA
 R ? I M V Y A P S R T A V V Y

NGAAGAATTTT
 NCTTCTTAAAA
 E E F

2531

Fig 9 SHEET 9

Fig 9
Sheet
10

45/75

CACCTCTCTGATGGCTCAGTAATTCCCGGAAACCAATTCAGTTA
GTGGAGAGACTACCGAGTCATTAAGGGCCTTTGGTTAAGTCAAT 2070
H L S D G S V I P G N Q F S Y

Nco I

AGATACCATGGGTTGCAAGAATTTGACCGGGCTATGCAGTATCT
TCTATGGTACCCAACGTTCTTAACTGGCCCGATACGTCATAGA 2160
R Y H G L Q E F D R A M Q Y L

CGAAAGGATGAAGGAGATAGGATGATTGTATTTGAAARAGGAAA
GCTTTCCTACTTCCTCTATCCTACTAACATAAACTTTTTCCTTT 2250
R K D E G D R M I V F E ? G N

TATCGCATAGGCTGCCTGAAGCCTGGAAAATACAAGGTTGGCTT
ATAGCGTATCCGACGGACTTCGGACCTTTTATGTTCCAACCGAA 2340
Y R I G C L K P G K Y K V G L

Ssp I

AATGCCGAATATTTACCTCTGAAGGATCGTATGATGATCGYCC
TTACGGCTTATAAAGTGGAGACTTCCTAGCATACTACTAGCRGG 2430
N A E Y F T S E G S Y D D R P

GCACTAGTAGACAAANTAGAAGNAGAAGAAGAAGAANCCGN
CGTGATCATCTGTTTATCTTCNTCTTCTTCTTCTTNGGCN 2520
A L V D K ? E ? E E E E ? ?

Fig. 9 SHEET 10

46/75

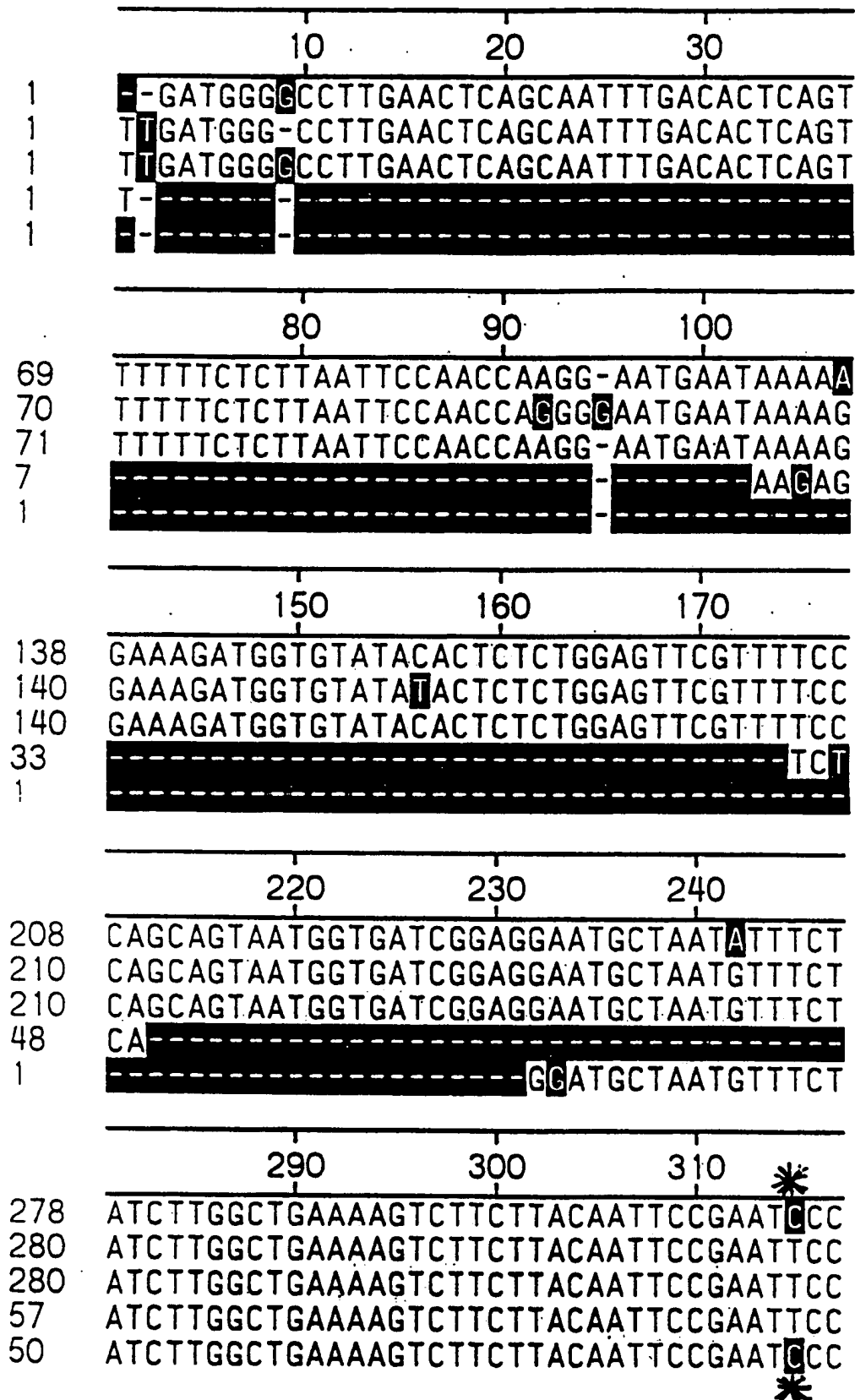
Fig.10
Sheet 2

Fig.10 SHEET 1

47/75

40	50	60	70	
TAGTTACACTGCGCATCACTTATCAGATCTCTAT				10con. seq
TAGTTACACTCCTATCACTTATCAGATCTCTAT				11con. seq
TAGTTACACTCCTATCACTTATCAGATCTCTAT				19con. seq
-----CATTA-----				86CON. SEQ
-----				pcrsbe2con. seq

110	120	130	140	
GATAGATTTGTAAAAACCCTAAGGAGAGAAGAA				10con. seq
GATAGATTTGTAAAAACCCTAAGGAGAGAAGAA				11con. seq
GATAGATTTGTAAAAACCCTAAGGAGAGAAGAA				19con. seq
GAGAAATT-----AACTATGAGAGGA-----				86CON. SEQ
-----				pcrsbe2con. seq

180	190	200	210	
TACTGTTCCATCAGTGTACAAATCTAATGGATT				10con. seq
TACTGTTCCATCAGTGTACAAATCTAATGGATT				11con. seq
TACTGTTCCATCAGTGTACAAATCTAATGGATT				19con. seq
CACCAT--CACCA-----T				86CON. SEQ
-----				pcrsbe2con. seq

250	260	270	280	
GTATTCTTGAAAAA A CACTCTCTTTCACGGAAG				10con. seq
GTATTCTTGAAAAAGCACTCTCTTTCACGGAAG				11con. seq
GTATTCTTGAAAAAGCACTCTCTTTCACGGAAG				19con. seq
-----CATGG--G				86CON. SEQ
GTATTCTTGAAAAAGCACTCTCTTTCACGGAAG				pcrsbe2con. seq

320	330	340	350	
GACCTTCTACA A TTGCAGCATCGGGGAAAGTCC				10con. seq
GACCTTCTACAGTTGCAGCATCGGGGAAAGTCC				11con. seq
GACCTTCTACAGTTGCAGCATCGGGGAAAGTCC				19con. seq
GACCTTCTACAGTTGCAGCATCGGGGAAAGTCC				86CON. SEQ
GACCTTCTACAGTTGCAGCATCGGGGAAAGTCC				pcrsbe2con. seq

Fig. 10 SHEET 2

48/75

	360	370	380
348	TTGTGCCTGGAAT	CCAGAGTGATAGCTCCTCATCCTC	
350	TTGTGCCTGGAACCCAGAGTGATAGCTCCTCATCCTC		
350	TTGTGCCTGGAACCCAGAGTGATAGCTCCTCATCCTC		
127	TTGTGCCTGGAACCCAGAGTGATAGCTCCTCATCCTC		
120	TTGTGCCTGGAAY	CCAGAGTGATAGCTCCTCATCCTC	
	430	440	450
418	AGAAAATTCCCCAGCATCAACTGATGTAGATAGTTCA		
420	AGAAAATTCCCCAGCATCAACTGATGTAGATAGTTCA		
420	AGAAAATTCCCCAGCATCAACTGATGTAGATAGTTCA		
197	AGAAAATTCCCCAGCATCAACTGATGTAGATAGTTCA		
190	AGAAAATTCCCCAGCATCAACTGATGTAGATAGTTCA		
	500	510	520
488	AACGATGACGTTGAGCCGTCAAGTGATCTTACAGGAA		
490	AACGATGACGTTGAGCCGTCAAGTGATCTTACAGGAA		
490	AACGATGACGTTGAGCCGTCAAGTGATCTTACAGGAA		
267	AACGATGACGTTGAGCCGTCAAGTGATCTTACAGGAA		
260	AACGATGACGTTGAGCCGTCAAGTGATCTTACAGGAA		
	570	580	590
558	AACTACAAGAAGGTGGTAAACTGGAGGAGTCTAAAAC		
560	AACTACAAGAAGGTGGTAAACTGGAGGAGTCTAAAAC		
560	AACTACAAGAAGGTGGTAAACTGGAGGAGTCTAAAAC		
337	AACTACAAGAAGGTGGTAAACTGGAGGAGTCTAAAAC		
330	AACTACAAGAAGGTGGTAAACTGGAGGAGTCTAAAAC		
	640	650	660
628	ATCTGATAGGATCAGAGAGAGGGGCATCCCTCCACCT		
630	ATCTGATAGGATCAGAGAGAGGGGCATCCCTCCACCT		
630	ATCTGATAGGATCAGAGAGAGGGGCATCCCTCCACCT		
407	ATCTGATAGGATCAGAGAGAGGGGCATCCCTCCACCT		
400	ATCTGATAGGATCAGAGAGAGGGGCATCCCTCCACCT		

Fig. 10
Sheet 4

Fig. 10 SHEET 3

49/75

390	400	410	420	
AACAGAT	CAATTTGAGTTC	CTGAGACATCTCC	10con. seq	
AACAGACCAATTTGAGTTC	ACTGAGACATCTCC	11con. seq		
AACAGACCAATTTGAGTTC	ACTGAGACATCTCC	19con. seq		
AACA	ACCAATTTGAGTTC	ACTGAGACATCTCC	86CON. SEQ	
AACAGACCAATTTGAGTTC	ACTGAGACATCTCC	pcrsbe2con. seq		
460	470	480	490	
ACAATGGAACACGCTAGCCAGATTAA	AACTGAG	10con. seq		
ACAATGGAACACGCTAGCCAGATTAA	AACTGAG	11con. seq		
ACAATGGAACACGCTAGCCAGATTAA	AACTGAG	19con. seq		
ACAATGGAACACGCTAGCCAGATTAA	AACTGAG	86CON. SEQ		
ACAATGGAACACGCTAGCCAGATTAA	AACTGAG	pcrsbe2con. seq		
530	540	550	560	
GTGTTGAAGAGCTGGATTTTGCTTCATCACTAC	10con. seq			
GTGTTGAAGAGCTGGATTTTGCTTCATCACTAC	11con. seq			
GTGTTGAAGAGCTGGATTTTGCTTCATCACTAC	19con. seq			
GTGTTGAAGAGCTGGATTTTGCTTCATCACTAC	86CON. SEQ			
GTGTTGAAGAGCTGGATTTTGCTTCATCACTAC	pcrsbe2con. seq			
600	610	620	630	
ATTAAATACTTCTGAAGAGACAATTATTGATGA	10con. seq			
ATTAAATACTTCTGAAGAGACAATTATTGATGA	11con. seq			
ATTAAATACTTCTGAAGAGACAATTATTGATGA	19con. seq			
ATTAAATACTTCTGAAGAGACAATTATTGATGA	86CON. SEQ			
ATTAAATACTTCTGAAGAGACAATTATTGATGA	pcrsbe2con. seq			
670	680	690	700	
GGACTTGGTCAGAAGATTTATGAAATAGACCCC	10con. seq			
GGACTTGGTCAGAAGATTTATGAAATAGACCCC	11con. seq			
GGACTTGGTCAGAAGATTTATGAAATAGACCCC	19con. seq			
GGACTTGGTCAGAAGATTTATGAAATAGACCCC	86CON. SEQ			
GGACTTGGTCAGAAGATTTATGAAATAGACCCC	pcrsbe2con. seq			

Fig.10 SHEET 4

50/75

	710	720	730
698	CTTTTGACAAACTATCGTCAACACCTTGATTACAGGT		
700	CTTTTGACAAACTATCGTCAACACCTTGATTACAGGT		
700	CTTTTGACAAACTATCGTCAACACCTTGATTACAGGT		
477	CTTTTGACAAACTATCGTCAACACCTTGATTACAGGT		
470	CTTTTGACAAACTATCGTCAACACCTTGATTACAGGT		
	780	790	800
768	ACAAGTATGAGGGTGGTTTGGGAAGCTTTTCTCGTGG		
770	ACAAGTATGAGGGTGGTTTGGGAAGC-TTTTCTCGTGG		
770	ACAAGTATGAGGGTGGTTTGGGAAGCCTTTTCTCGTGG		
547	ACAAGTATGAGGGTGGTTTGGGAAGCTTTTCTCGTGG		
540	ACAAGTATGAGGGTGGTTTGGGAAGCTTTTCTCGTGG		
	850	860	870
838	AGGTATCACTTACCGTGAGTGGGCTCCTGGTGCCCAG		
839	AGGTATCACTTACCGTGAGTGGGCTCCTGGTGCCCAG		
840	AGGTATCACTTACCGTGAGTGGGCTCCTGGTGCCCAG		
617	AGGTATCACTTACCGTGAGTGGGCTCCTGGTGCCCAG		
610	AGGTATCACTTACCGTGAGTGGGCTCCTGGTGCCCAG		
	920	930	940
908	GACGCAAATGCTGACATTATGACTCGGAATGAATTTG		
909	GACGCAAATGCTGACATTATGACTCGGAATGAATTTG		
910	GACGCAAATGCTGACATTATGACTCGGAATGAATTTG		
687	GACGCAAATGCTGACATTATGACTCGGAATGAATTTG		
680	GACGCAAATGCTGACATTATGACTCGGAATGAATTTG		
	990	1000	1010
978	ATGGTTCTCCTGCAATTCCTCATGGGTCCAGAGTGAA		
979	ATGGTTCTCCTGCAATTCCTCATGGGTCCAGAGTGAA		
980	ATGGTTCTCCTGCAATTCCTCATGGGTCCAGAGTGAA		
757	ATGGTTCTCCTGCAATTCCTCATGGGTCCAGAGTGAA		
750	ATGGTTCTCCTGCAATTCCTCATGGGTCCAGAGTGAA		

Fig.10
Sheet 6

Fig.10 SHEET 5

51/75

740	750	760	770	
ATTACAGTACAAGAACTGAGGGAGGCAATTG				10con. seq
ATTACAGTACAAGAACTGAGGGAGGCAATTG				11con. seq
ATTACAGTACAAGAACTGAGGGAGGCAATTG				19con. seq
ATTACAGTACAAGAACTGAGGGAGGCAATTG				86CON. SEQ
ATTACAGTACAAGAACTGAGGGAGGCAATTG				pcrsbe2con. seq
810	820	830	840	
TTATGAAA C AATGGGTTTCACTCGTAGTGCTAC				10con. seq
TTATGAAAAAATGGGTTTCACTCGTAGTGCTAC				11con. seq
TTATGAAAAAATGGGTTTCACTCGTAGTGCTAC				19con. seq
TTATGAAAAAATGGGTTTCACTCGTAGTGCTAC				86CON. SEQ
TTATGAAAAAATGGGTTTCACTCGTAGTGCTAC				pcrsbe2con. seq
880	890	900	910	
TCAGCTGCCCTCATTGG G GATTTCAACAATTGG				10con. seq
TCAGCTGCCCTCATTGGAGATTTCAACAATTGG				11con. seq
TCAGCTGCCCTCATTGGAGATTTCAACAATTGG				19con. seq
TCAGCTGCCCTCATTGGAGATTTCAACAATTGG				86CON. SEQ
TCAGCTGCCCTCATTGGAGATTTCAACAATTGG				pcrsbe2con. seq
950	960	970	980	
GTGTCTG A GAGATTTTTCTGCCAAATAATGTGG				10con. seq
GTGTCTGGGAGATTTTTCTGCCAAATAATGTGG				11con. seq
GTGTCTGGGAGATTTTTCTGCCAAATAATGTGG				19con. seq
GTGTCTGGGAGATTTTTCTGCCAAATAATGTGG				86CON. SEQ
GTGTCTGGGAGATTTTTCTGCCAAATAATGTGG				pcrsbe2con. seq
1020	1030	1040	1050	
GATACGTATGGACACTCCATCAGGTGTTAAGGA				10con. seq
GATACGTATGGACACTCCATCAGGTGTTAAGGA				11con. seq
GATACGTATGGACACTCCATCAGGTGTTAAGGA				19con. seq
GATACGTATGGACACTCCATCAGGTGTTAAGGA				86CON. SEQ
GATACG Y ATGGACACTCCATCAGGTGTTAAGGA				pcrsbe2con. seq

Fig. 10 SHEET 6

52/75

	1060	1070	1080
1048	TTCCATTCTGCTTGGATCAACTACTCTTTACAGCTT		
1049	TTCCATTCTGCTTGGATCAACTACTCTTTACAGCTT		
1050	TTCCATTCTGCTTGGATCAACTACTCTTTACAGCTT		
827	TTCCATTCTGCTTGGATCAACTACTC--TACAGCTT		
820	TTCCATTCTGCTTGGATCAACTACTCTTTACAGCTT		

	1130	1140	1150
1118	GATCCACCCGAAGAGGAGAGGTATATCTTCCAACACC		
1119	GATCCACCCGAAGAGGAGAGGTATATCTTCCAACACC		
1120	GATCCACCCGAAGAGGAGAGGTATATCTTCCAACACC		
895	GATCCACCCGAAGAGGAGAGGTATATCTTCCAACACC		
890	GATCCACCCGAAGAGGAGAGGTATRTCTTCCAACACC		

	1200	1210	1220
1188	ATGAATCTCATATTGGAATGAGTAGTCCGGAGCCTAA		
1189	ATGAATCTCATATTGGAATGAGTAGTCCGGAGCCTAA		
1190	ATGAATCTCATATTGGAATGAGTAGTCCGGAGCCTAA		
965	ATGAATCTCATATTGGAATGAGTAGTCCGGAGCCTAA		
960	ATGAATCTCATATTGGAATGAGTAGTCCGGAGCCTAA		

	1270	1280	1290	*
1258	TCTTCCTCGCATAAAAAAAGCTTGGGTACAATGCGGT			
1259	TCTTCCTCGCATAAAAAA-GCTTGGGTACAATGCGCT			
1260	TCTTCCTCGCATAAAAAA-GCTTGGGTACAATGCGCT			
1035	TCTTCCTCGCATAAAAAA-GCTTGGGTACAATGCGCT			
1030	TCTTCCTCGCATAAAAAA-SCTTGGGTACAATGCGGT			

	1340	1350	1360
1328	TGCTAGTTTTGGTTATCATGTCACAAATTTTTTTGCA		
1328	TGCTAGTTTTGGTTATCATGTCACAAATTTTTTTGCA		
1329	CGCTAGTTTTGGTTATCATGTCACAAATTTTTTTGCA		
1104	TGCTAGTTTTGGTTATCATGTCACAAATTTTTTTGCA		
1099	TGCTAGTTTTGGTTATCATGTCACAAATTTTTTTGCA		

Fig.10
Sheet 8

Fig.10 SHEET 7

53/75

1090	1100	1110	1120	
CCTGATGAAATTCCATATAATGGAATATATTAT	10con. seq			
CCTGATGAAATTCCATATAATGGAATATATTAT	11con. seq			
CCTGATGAAATTCCATATAATGGAATACATTAT	19con. seq			
CCTGATGAAATTCCATATAATGGAATATATTAT	86CON. SEQ			
CCTGATGAAATTCCATATAATGGAATATATTAT	pcrsbe2con. seq			
1160	1170	1180	1190	
CACGGCCAAAGAAACCAAAGTCGCTGAGAATAT	10con. seq			
CACGGCCAAAGAAACCAAAGTCGCTGAGAATAT	11con. seq			
CACGGCCAAAGAAACCAAAGTCGCTGAGAATAT	19con. seq			
CACGGCCAAAGAAACCAAAGTCGCTGAGAATAT	86CON. SEQ			
CACGGCCAAAGAAACCAAAGTCGCTGAGAATAT	pcrsbe2con. seq			
1230	1240	1250	1260	
AATTAACTCATACGTGAATTTTAGAGATGAAGT	10con. seq			
AATTAACTCATACGTGAATTTTAGAGATGAAGT	11con. seq			
AATTAACTCATACGTGAATTTTAGAGATGAAGT	19con. seq			
AATTAACTCATACGTGAATTTTAGAGATGAAGT	86CON. SEQ			
AATTAACTCATACGTGAATTTTAGAGATGAAGT	pcrsbe2con. seq			
1300	1310	1320	1330	
GCAAATTATGGCTATTCAAGAGCATTCTTATTA	10con. seq			
GCGAATTATGGCTATTCAAGAGCATTCTTATTA	11con. seq			
GCAAATTATGGCTATTCAAGAGCATTCTTATTA	19con. seq			
GCAAATTATGGCTATTCAAGAGCATTCTTATTA	86CON. SEQ			
GCAAATTATGGCTATTCAAGAGCATTCTTATTA	pcrsbe2con. seq			
1370	1380	1390	1400	
CCAAGCAGCCGTTTTGGAACGCCCGACGACCTT	10con. seq			
CCAAGCAGCCGTTTTGGAACGCCCGACGACCTT	11con. seq			
CCAAGCAGCCGTTTTGGAACGCCCGACGACCTT	19con. seq			
CCAAGCAGCCGTTTTGGAACGCCCGACGACCTT	86CON. SEQ			
CCAAGCAGCCGTTTTGGAACGCCCGACGACCTT	pcrsbe2con. seq			

Fig. 10 SHEET 8

54/75

	1410	1420	1430
1398	AAGTCTTTGATTGATAAAGCTCATGAGCTAGGAATTG		
1398	AAGTCTTTGATTGATAAAGCTCATGAGCTAGGAATTG		
1399	AAGTCTTTGATTGATAAAGCTCATGAGCTAGGAATTG		
1174	AAGTCTTTGATTGATAAAGCTCATGAGCTAGGAATTG		
1169	AAGTCTTTGATTGATAAAGCTCATGAGCTAGGAATTG		
	1480	1490	1500
1468	CAAATAAATACTTTAGATGGACTGAACATGTTTGACGG		
1468	CAAATAAATACTTTAGATGGACTGAACATGTTTGACGG		
1469	CAAATAAATACTTTAGATGGACTGAACATGTTTGAC	TG	
1244	CAAATAAATACTTTAGATGGACTGAACATGTTTGACGG		
1239	CAAATAAATACTTTAGATGGACTGAACATGTTTGACGG		
	1550	1560	1570
1538	TGGTTATCATTGGATGTGGGATT	TCCGCCTCTTTAAC	
1538	TGGTTATCATTGGATGTGGGATT	TCCGCCTCTTTAAC	
1539	TGGTTATCATTGGATGTGGGATT	TCCGCCTCTTTAAC	
1314	TGGTTATCATTGGATGTGGGATT	TCCGCCTTTTAACT	
1309	TGGTTATCATTGGATGTGGGATT	TCCGCCTCTTTAAC	
	1620	1630	1640
1608	TCAAATGCGAGATGGTGGTTGGATGAGTTCAAATTTG		
1607	TCAAATGCGAGATGGTGGTTGGATGAGTTCAAATTTG		
1609	TCAAATGCGAGATGGTGGTTGGATG	CGTTCAAATTTG	
1384	TCAAATGCGAGATGGTGGTTGGATGAGTTCAAATTTG		
1379	TCAAATGCGAGATGGTGGTTGGATGAGTTCAAATTTG		
	1690	1700	1710
1678	TGTGTA	CTCACCACGGATTATCGGTGGGATTCACTGG	
1677	TGTATA	CTCACCACGGATTATCGGTGGGATTCACTGG	
1679	TGTATA	CTCACCACGGATTATCGGTGGGATTCACTGG	
1454	TGTATA	CTCACCACGGATTATCGGTGGGATTCACTGG	
1449	TGTATA	CTCACCACGGATTATCGGTGGGATTCACTGG	

Fig. 10
Sheet 10

55/75

1440	1450	1460	1470	
TTGTTCTCATGGACATTGTT	CACAGCCATGCAT	10con. seq		
TTGTTCTCATGGACAT	GTTTACAGCCATGCAT	11con. seq		
TTGTTCTCATGGACATTGTT	CACAGCCATGCAT	19con. seq		
TTGTTCTCATGGACATTGTT	CACAGCCATGCAT	86CON. SEQ		
TTGTTCTCATGGACATTGTT	CACAGCCATGCAT	pcrsbe2con. seq		
1510	1520	1530	1540	
CACAGATAGTTGTTACTTT	CACTCTGGAGCTCG	10con. seq		
CACCGATAGTTGTTACTTT	CACTCTGGAGCTCG	11con. seq		
CACCGATAGTTGTTACTTT	CACTCTGGAGCTCG	19con. seq		
CACCGATAGTTGTTACTTT	CACTCTGGAGCTCG	86CON. SEQ		
CACAGATAGTTGTTACTTT	CACTCTGGAGCTCG	pcrsbe2con. seq		
1580	1590	1600	1610	
TATGGAAACTGGGAGGTACT	TAGGTATCTTCTC	10con. seq		
TATGGAAACTGGGAGGTACT	TAGGTATCTTCTC	11con. seq		
TATGGAAACTGGGAGGTACT	TAGGTATCTTCTC	19con. seq		
TATGGAAACTGGGAGGTACT	TAGGTATCTTCTC	86CON. SEQ		
TATGGAAACTGGGAGGTACT	TAGGTATCTTCTC	pcrsbe2con. seq		
1650	1660	1670	1680	
ATGGATTTAGATTTGATGGT	GTGACATCAATGA	10con. seq		
ATGGATTTAGATT	GATGGTGTGACATCAATGA	11con. seq		
ATGGATTTAGATTTGATGGT	GTGACATCAATGA	19con. seq		
ATGGATTTAGATTTGATGGT	GTGACATCAATGA	86CON. SEQ		
ATGGATTTAGATTTGATGGT	GTGACATCAATGA	pcrsbe2con. seq		
1720	1730	1740	1750	
GAACTACGAGGAATACTTT	GGACTCGCAACTGA	10con. seq		
GAACTACGAGGAATACTTT	GGACTCGCAACTGA	11con. seq		
GAACTACGAGGAATACTTT	GGACTCGCAACTGA	19con. seq		
GAACTACGAGGAATACTTT	GGACTCGCAACTGA	86CON. SEQ		
GAACTACGAGGAATACTTT	GGACTCGCAACTGA	pcrsbe2con. seq		

Fig. 10 SHEET 10

56/75

	1760	1770	1780
1748	TGTGGATGCTGTTGTGTATCTGATGCTGGTCAACGAT		
1747	TGTGGATGCTGTTGTGTATCTGATGCTGGTCAACGAT		
1749	TGTGGATGCTGTTGTGTATCTGATGCTGGTCAACGAT		
1524	TGTGGATGCTGTTGTGTATCTGATGCTGGTCAACGAT		
1519	TGTGGATGCTGTTGTGTATCTGATGCTGGTCAACGAT		
	1830	1840	1850
1818	ATTGGTGAAGATGTTAGCGGAATGCCGACATTTTGT		
1817	ATTGGTGAAGATGTTAGCGGAATGCCGACATTTTGT		
1819	ATTGGTGAAGATGTTAGCGGAATGCCGACATTTTGT		
1594	ATTGGTGAAGATGTTAGCGGAATGCCGACATTTTGT		
1589	ATTGGTGAAGATGTTAGCGGAATGCCGACATTTTGT		
	1900	1910	1920
1888	ATCGGCTGCATATGGCAATTGCTGATAAATGGATTGA		
1887	ATCGGCTGCATATGGCAATTGCTGATAAATGGATTGA		
1889	ATCGGCTGCATATGGCAATTGCTGATAAATGGATTGA		
1664	ATCGGCTGCATATGGCAATTGCTGATAAATGGATTGA		
1659	ATCGGCTGCATATGGCAATTGCTGATAAATGGATTGA		
	1970	1980	1990
1958	GGGTGATATTGTTTCATACACTGACAAATAGAAGATGG		
1957	GGGTGATATTGTTTCATACACTGACAAATAGAAGATGG		
1959	GGGTGATATTGTTTCATACACTGACAAATAGAAGATGG		
1734	GGGTGATATTGTTTCATACACTGACAAATAGAAGATGG		
1729	GGGTGATATTGTTTCATACACTGACAAATAGAAGATGG		
	2040	2050	2060
2028	GATCAAGCTCTAGTCGGTGATAAACTATAGCATTCT		
2027	GATCAAGCTCTAGTCGGTGATAAACTATAGCATTCT		
2029	GATCAAGCTCTAGTCGGTGATAAACTATAGCATTCT		
1804	GATCAAGCTCTAGTCGGTGATAAACTATAGCATTCT		
1799	GATCAAGCTCTAGTCGGTGATAAACTATAGCATCT		

Fig.10
Sheet 12

57/75

1790	1800	1810	1820	
CTTATTTCATGGGCTTTTCCCAGATGCAATTACC	10con. seq			
CTTATTTCATAGGCTTTTCCCAGATGCAATTACC	11con. seq			
CTTATTTCATGGGCTTTTCCCAGATGCAATTACC	19con. seq			
CTTATTTCATGGGCTTTTCCCAGATGCAATTACC	86CON. SEQ			
CTTATTTCACGGGCTTTTCCCAGATGCAATTACC	pcrsbe2con. seq			
1860	1870	1880	1890	
TTCCCGTTCAAGATGGGGGTGTTGGCTTTGACT	10con. seq			
TTCCCGTTCAAGATGGGGGTGTTGGCTTTGACT	11con. seq			
TTCCCGTCAAGAGGGGGGTGTTGGCTTTGACT	19con. seq			
TTCCCGTTCAAGATGGGGGTGTTGGCTTTGACT	86CON. SEQ			
TTCCCGTTCAAGATGGGGGTGTTGGCTTTGACT	pcrsbe2con. seq			
1930	1940	1950	1960	
GTTGCTCAAGAAACGGGATGAGGATTGGAGAGT	10con. seq			
GTTGCTCAAGAAACGGGATGAGGATTGGAGAGT	11con. seq			
GTTGCTCAAGAAACGGGATGAGGATTGGAGAGT	19con. seq			
GTTGCTCAAGAAACGGGATGAGGATTGGAGAGT	86CON. SEQ			
GTTGCTCAAGAAACGGGATGAGGATTGGAGAGT	pcrsbe2con. seq			
2000	2010	2020	2030	
TCGGAAAAGTGTGTTTCATACGCTGAAAGTCAT	10con. seq			
TCGGAAAAGTGTGTTTCATACGCTGAAAGTCAT	11con. seq			
TCGGAAAAGTGTGTTTCATACGCTGAAAGTCAT	19con. seq			
TCGGAAAAGTGTGTTTCATACGCTGAAAGTCAT	86CON. SEQ			
TCGGAAAAGTGTGTTTCATMCGCTGAAAGTCAT	pcrsbe2con. seq			
2070	2080	2090	2100	
GGCTGATGGACAAGGATATGTATGATTTTATGG	10con. seq			
GGCTGATGGACAAGGATATGTATGATTTTATGG	11con. seq			
GGCTGATGGACAAGGATATGTATGATTTTATGG	19con. seq			
GGCTGATGGACAAGGATATGTATGATTTTATGG	86CON. SEQ			
GGCTGATGGACAAGGATATGTATGATTTTATGG	pcrsbe2con. seq			

Fig. 10 SHEET 12

58/75

	2110	*	2120	2130
2098	CTCTGGATAGACCGT	CAACATCATT	AATAGATCGTGG	
2097	CTCTGGATAGACCG	CAACATCATT	AATAGATCGTGG	
2099	CTCTGGATAGACCGT	CAACATCATT	AATAGATCGTGG	
1874	CTCTGGATAGACCG	CAACATCATT	AATAGATCGTGG	
1869	CTCTGGATAGACCGY	CAACAY	CATT	AATAGATCGTGG
	2180		2190	2200
2168	TATGGGATTAGGAGGAGAAGGGTACCTAAATTT	CATG		
2167	TATGGGATTAGGAGGAGAAGGGTACCTAAATTT	CATG		
2169	TATGGGATTAGGAGGAGAAGGGTACCTAAATTT	CATG		
1944	TATGGGATTAGGAGGAGAAGGGTACCTAAATTT	CATG		
1939	TATGGGATTAGGAGGAGAAGGGTACCTAAATTT	CATG		
	2250	*	2260	2270
2238	TTCCCTAGGGCTGAACAACACCTCTCTGATGGCTCAG			
2237	TTCCCTAGGGCTGA	CCACACCT	TTCTGATGGCTCAG	
2239	TTCCCTAGGGCTGAACAACACCTCTCTGATGGCTCAG			
2014	TTCCCTAGGGCTGAACAACACCTCTCTGATG	ACTCAG		
2009	TTCCCTAGGGCTGAR	CAACACCTCTCTGATGGCTCAG		
	2320		2330	2340
2308	GCAGACGGAGATTTGACCTGGGAGATGCAGAATATTT			
2307	GCAGACGGAGATTTGACCTGGGAGATGCAGAATATTT			
2309	GCAGACGGAGATTTGACCTGGGAGATGCAGAATATTT			
2084	GCAGACGGAGATTTGACCTGGGAGATGCAGAATATTT			
2079	GCAGACGGAGATTTGACCTGGGAGATGCAGAATATTT			
	2390		2400	2410
2378	TATGCAGTATCTTGAAGATAAATATGAGTTT	TATGACT		
2377	TATGCAGTATCTTGAAGATAAATATGAGTTT	TATGACT		
2379	TATGCAGTATCTTGAAGATAAATATGAGTTT	TATGACT		
2154	TATGCAGTATCTTGAAGATAAATATGAGTTT	TATGACT		
2149	TATGCAGTATCTTGAAGATAAATATGAGTTT	TATGACT		

Fig.10
Sheet 14

Fig. 10 SHEET 13

59/75

2140	2150	2160	2170	
GATAGCATTACACAAGATGATTAGGCTTGTAAC				10con. seq
GATAGCATTGCACAAGATGATTAGGCTTGTAAC				11con. seq
GATAGCATTGCACAAGATGATTAGGCTTGTAAC				19con. seq
GATAGCATTGCACAAGATGATTAGGCTTGTAAC				86CON. SEQ
GATAGCATTGCACAAGATGATTAGGCTTGTAAC				pcrsbe2con. seq

2210	2220	2230	2240	
GGAAATGAATTCGGCCACCCTGAGTGGATTGAT				10con. seq
GGAAATGAATTCGGCCACCCTGAGTGGATTGAT				11con. seq
GGAAATGAATTCGGCCACCCTGAGTGGATTGAT				19con. seq
GGAAATGAATTCGGCCACCCTGAGTGGATTGAT				86CON. SEQ
GGAAATGAATTCGGCCACCCTGAGTGGATTGAT				pcrsbe2con. seq

2280	2290	2300	2310	
TAATTCCCAGAAACCAATTCAGTTATGATAAAT				10con. seq
TAATTCCCGGAAACCAATTCAGTTATGATAAAT				11con. seq
TAATCCCGGAAACCAATTCAGTTATGATAAAT				19con. seq
TAATTCCCGGAAACCAATTCAGTTATGATAAAT				86CON. SEQ
TAATTCCCGGAAACCAATTCAGTTATGATAAAT				pcrsbe2con. seq

2350	2360	2370	2380	
AAGATACCGTGGGTTGCAAGAATTTGACCGGGC				10con. seq
AAGATACCATGGGTTCAAGAATTTGACGGGC				11con. seq
AAGATACCGTGGGTTGCAAGAATTTGACCGGGC				19con. seq
AAGATACCGTGGGTTGCAAGAATTTGACCGGGC				86CON. SEQ
AAGATACCATGGGTTGCAAGAATTTGACCGGGC				pcrsbe2con. seq

2420	2430	2440	2450	
TCAGAACACCAGTTCATATCACGAAAGGATGAA				10con. seq
TCAGAACACCAGTTCATATCACGAAAGGATGAA				11con. seq
TCAGAACACCAGTTCATATCACGAAAGGATGAA				19con. seq
TCAGAACACCAGTTCATATCACGAAAGGATGAA				86CON. SEQ
TCAGAACACCAGTTCATATCACGAAAGGATGAA				pcrsbe2con. seq

Fig. 10 SHEET 14

60/75

	2460	2470	*	2480
2448	GGAGATAGGATGATTGTATTTGAAAAAGGAAACCTAG			
2447	GGAGATAGGATGATTGTATTTGAAAAGGAAACCTAG			
2449	GGAGATAGGATGATTGTATTTGAAAAAGGAAACCTAG			
2224	GGAGATAGGATGATTGTATTTGAAAAAGGAAACCTAG			
2219	GGAGATAGGATGATTGTATTTGAAAAGGAAACCTAG			
			*	
	2530	2540		2550
2518	ATTCAGACTATCGCATAGGCTGCCTGAAGCCTGGAAA			
2517	ATTCAGACTATCGCATAGGCTGCCTGAAGCCTGGAAA			
2519	ATTCAGACTATCGCATAGCTGCCTGAAGCCTGGAAA			
2294	ATTCAGACTATCGCATAGGCTGCCTGAAGCCTGGAAA			
2289	ATTCAGACTATCGCATAGGCTGCCTGAAGCCTGGAAA			
	2600	2610		2620
2588	TTTTGGTGGCTTCGGGAGAATTGATCATAATGCCGAA			
2587	TTTTGGTGGCTTCGGGAGAATTGATCATAATGCCGAA			
2589	TTTTGGTGGCTTCGGGAGAATTGATCATAATGCCGAA			
2364	TTTTGGTGGCTTCGGGAGAATTGATCATAATGCCGAA			
2359	TTTTGGTGGCTTCGGGAGAATTGATCATAATGCCGAA			
	2670	2680	*	2690
2658	CCTCGTTCAATTATGGTGTATGCACCTAGTAGAACAG			
2657	CCTCGTTCAATTATGGTGTATGCACCTAGTAGAACAG			
2659	CCTCGTTCAATTATGGTGTATGCACCTGTAGAACAG			
2434	CCTCGTTCAATTATGGTGTATGCACCTGTAGAACAG			
2429	CCTCGTTCAATTATGGTGTATGCACCTAGTAGAACAG			
			*	
	2740	2750		2760
2722	-----AAGAAGAAGAAGAAGAAGTAGCAGTAGT			
2722	-----AGAAAGTAGCAGTAGT			
2729	AAGAAGAAGAAGAAGAAGAAGAAGTAGCAGTAGT			
2501	AAGAAGAAGAAGAAGAAGAAGAAGTAGCAGTAGT			
2499	NAGAAGAAGAAGAAGAN-----			

Fig. 10
Sheet 16

Fig. 10 SHEET 15

61/75

2490	2500	2510	*	2520	
TTTTTGTCTTTAATTTTCACTGGACAAAAGCT					10con. seq
TTTTT C GTCTTTAATTTTCACTGGACAAATAGCT					11con. seq
TTTTTGTCTTTAATTTTCACTGGACAAAAGCT					19con. seq
TTTTTGTCTTTAATTTTCACTGGACAAAAGCT					86CON. SEQ
TTTTTGTCTTTAATTTTCACTGGACAAATAGCT					pcrsbe2con. seq
					*
2560	2570	2580		2590	
ATACAAGGTTGCCTTGGACTCAGATGATCCACT					10con. seq
ATACAAGGTTG T CTTGGACTCAGATGATCCACT					11con. seq
ATACAAGGTTGCCTTGGACTCAGATGATCCACT					19con. seq
ATACAAGGTTGCCTTGGACTCAGATGATCCACT					86CON. SEQ
ATACAAGGTTG C CTTGGACTCAGATGATCCACT					pcrsbe2con. seq
2630	*	2640	*	2650	2660
TATTTACCTTTGAAGGATGGTATGATGATCGT					10con. seq
TATTTACCT C TGAAGGAT C GTATGATGATCGT					11con. seq
TATTTACCTTTGAAGGATGGTATGATGATCGT					19con. seq
TATTTACCTTTGAAGGATGGTATGATGATCGT					86CON. SEQ
TATTTACCT C TGAAGGAT C GTATGATGATCGT					pcrsbe2con. seq
					*
					*
2700	2710	2720		2730	
CAGTGGTCTATGCACTAGTAGACAAAG----					10con. seq
CAGTGGTCTATGCACTAGTAGACAAACT----					11con. seq
CAGTGGTCTATGCACTAGTAGACAAAGAAGAAG					19con. seq
CAGTGGTCTATGCACTAGTAGACAAAG--AAG					86CON. SEQ
CAGTGGTCTATGCACTAGTAGACAAANTAGAAG					pcrsbe2con. seq
2770	2780	2790		2800	
AGAAGAAGTAGTAGTAGAAGAAGAATGAACGAA					10con. seq
AGAAGAA CCCAT T G -----AAGAATGAACGAA					11con. seq
AGAAGAAGTAGTAGTAGAAGAAGAATGAACGAA					19con. seq
AGAAGAAGTAGTAGTAGAAGAAGAATGAACGAA					86CON. SEQ
----- CCGNN GAAGAAT-----					pcrsbe2con. seq

Fig. 10 SHEET 16

62/75

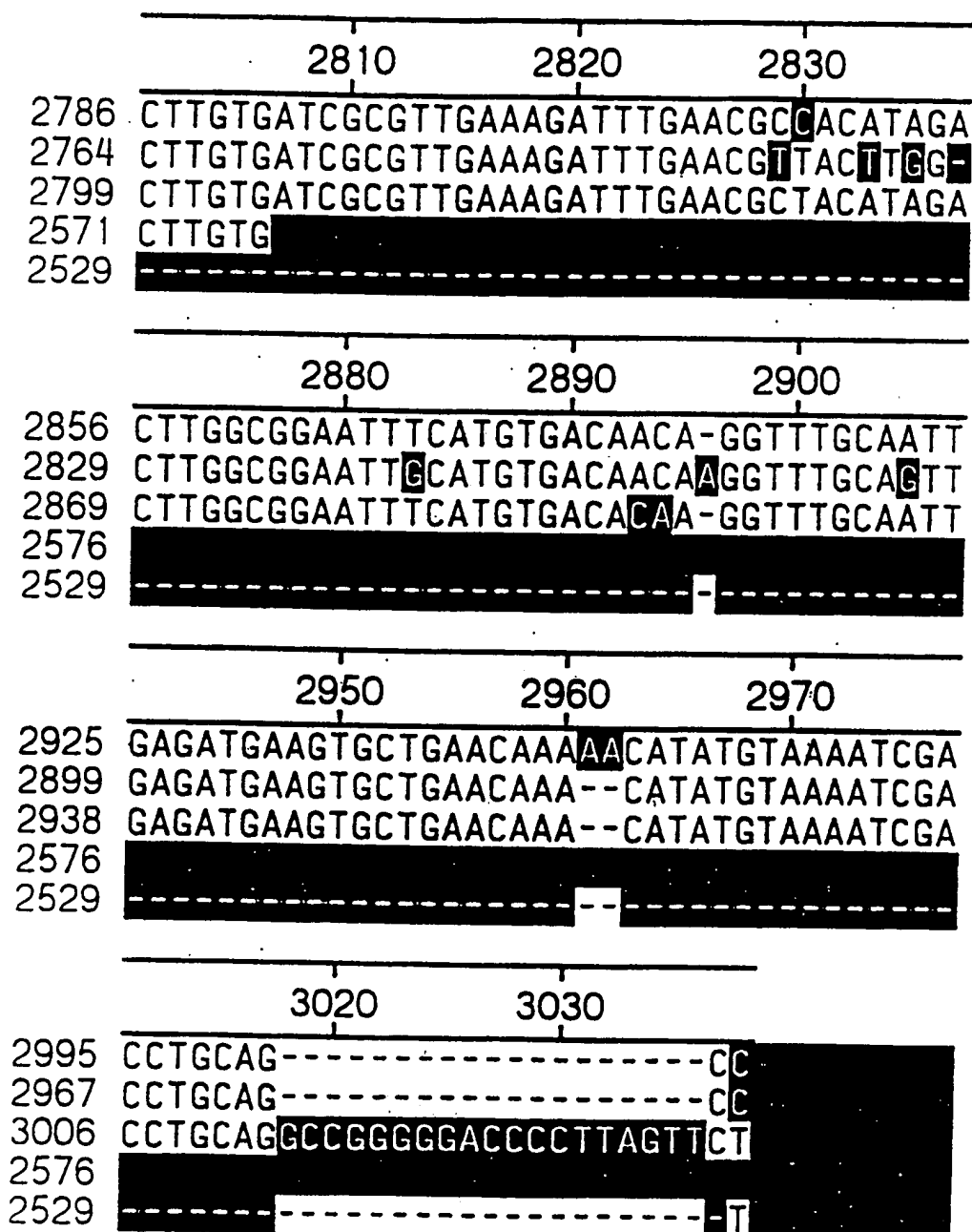
Fig. 10
Sheet 18

Fig. 10 SHEET 17

63/75




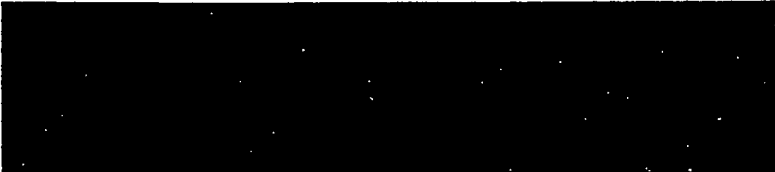
2840	2850	2860	2870	
GCTTCTTGACGTATCTGGCAATATTGCATTAGT				10con. seq
--TCATCCACATA--GAGCCTTCTTGACATCAGT				11con. seq
GCTTCTTGACGTATCTGGCAATATTGCATCAGT				19con. seq
				86CON. SEQ
-----				pcrsbe2con. seq
2910	2920	2930	2940	
CTTTCCACTATTAGTAGTGCAACGATATACGCA				10con. seq
CTTTCCACTATTAGTAGTCCACCGATATACGCA				11con. seq
CTTTCCACTATTAGTAGTGCAACGATATACGCA				19con. seq
				86CON. SEQ
-----				pcrsbe2con. seq
2980	2990	3000	3010	
TGAATTTATGTCTGAATGCTGGGACGATCGAATT				10con. seq
TGAATTTATGTCTGAATGCTGGGACGATCGAATT				11con. seq
TGAATTTATGTCTGAATGCTGGGACGATCGAATT				19con. seq
				86CON. SEQ
-----				pcrsbe2con. seq
				10con. seq
				11con. seq
				19con. seq
				86CON. SEQ
				pcrsbe2con. seq

Fig. 10 SHEET 18

64/75

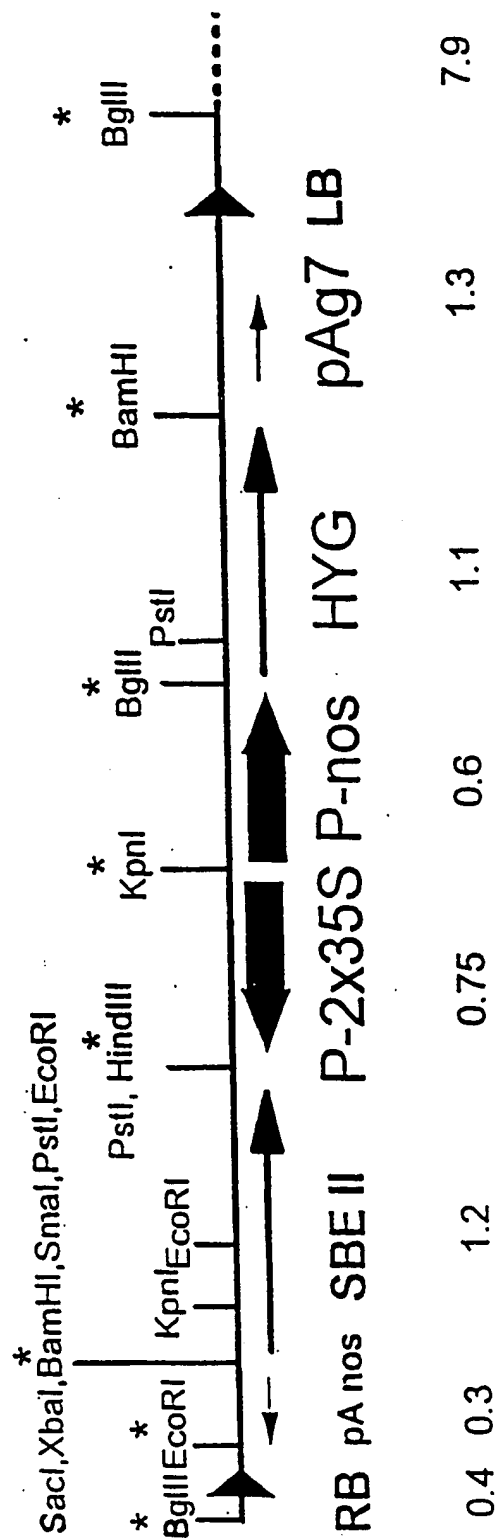
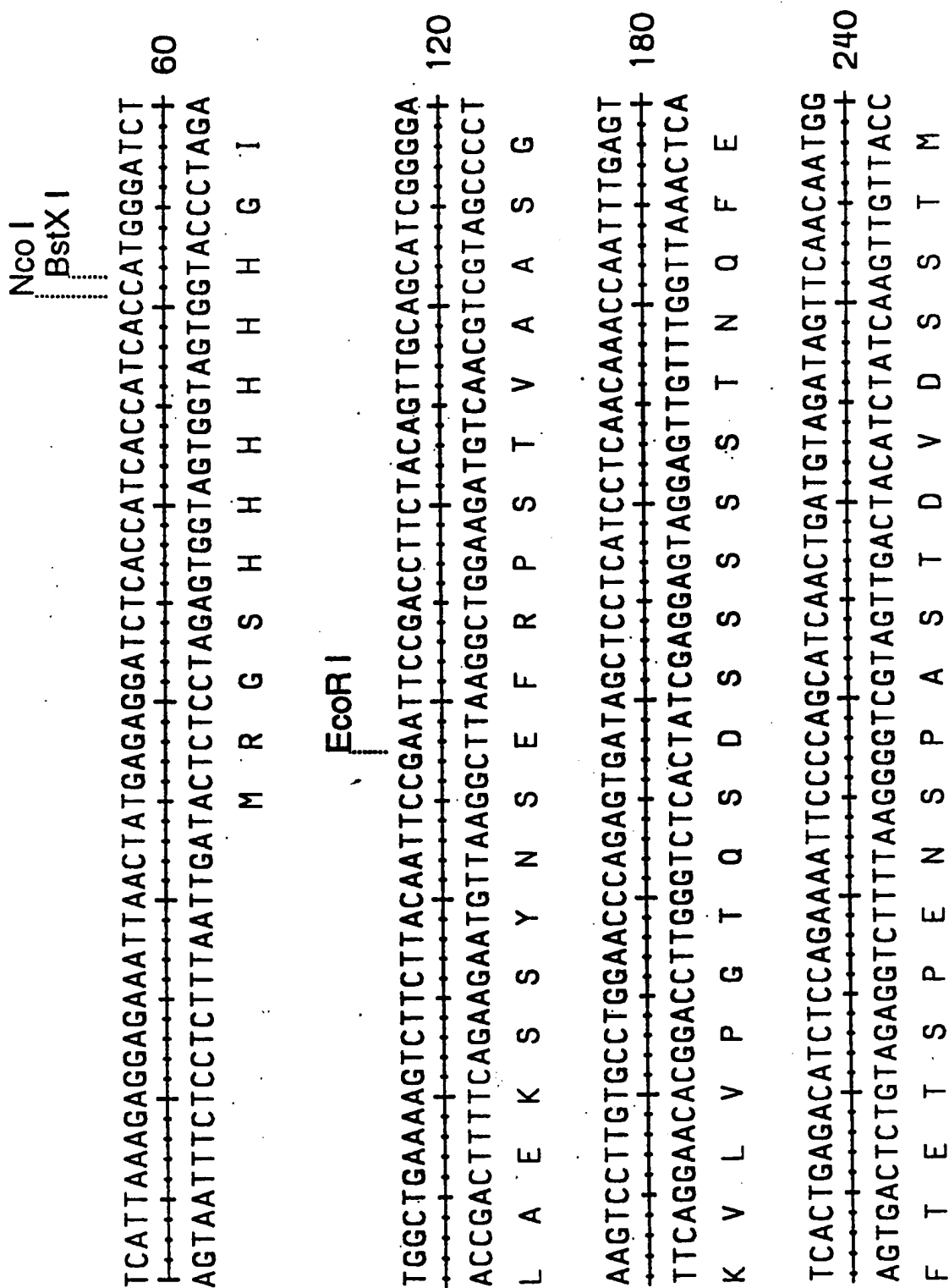


Fig. 11

65/75

Fig.12
SHEET 1

66/75

AACACGCTAGCCAGATTAAACTGAGAACGATGACGTTGAGCCGTCAAGTGATCTTACAG 300
TTGTGCGATCGGTCTAATTTTGACTCTTGCTACTGCAACTCGGCAGTTCACTAGAAATGTC
E H A S Q I K T E N D D V E P S S D L T

GAAGTGTGAAGAGCTGGATTTTGCTTCATCACTACAAGAGGTGGTAAACTGG 360
CTTCACAACCTTCTCGACCTAAACGAAGTAGTGATGTTGATGTTCTTCCACCATTTGACC
G S V E E L D F A S S L Q L Q E G G K L

AGGAGCTAAACATTAACTTCTGAAGAGACAATTATGATGAATCTGATAGGATCA 420
TCCTCAGATTTTGTAATTTATGAAGACTTCTCTGTAAATACTACTTAGACTATCCTAGT
E E S K T L N T S E E T I I D E S D R I

GAGAGGGGCATCCCTCCACCTGGACTTGGTCAGAAGATTTATGAAATAGACCCCTTT 480
CTCTCTCCCCGTAGGAGGTGGACCTGAACCAGTCTTCTAAATACTTTATCTGGGGGAAA
R E R G I P P P G L G Q K I Y E I D P L

Hinc II

TGACAAACTATCGTCAACACCTTGATTACAGGTATTCACAGTACAAGAACTGAGGGAGG 540
ACTGTTTGATAGCAGTTGTGGAACTAATGTCCATAAGTGTCATGTCTTTGACTCCCTCC
L T N Y R Q H L D Y R Y S Q Y K K L R E

Fig.12
SHEET 2

67/75

Fig 12
SHEET 3

Hind III

CAATTGACAAGTATGAGGGTGGTTTGGAAAGCTTTTCTCGTGGTTATGAAAAAATGGGTT
GTTAACTGTTTACATCTCCACCAACCTTCGAAAAGAGCACCAATACTTTTTTACCCAA
A I D K Y E G G L E A F S R G Y E K M G 600

Pvu II

TCACTCGTAGTGCTACAGGTATCACTTACCGTGAGTGGCTCCTGGTGCCAGTCAGCTG
AGTGAGCATCACGATGTCCATAGTGAATGGCACTCACCCGAGGACCACGGTTCAGTCGAC
F T R S A T G I T Y R E W A P G A O S A 660

CCCTCATTGGAGATTTCACAATGGGACGCAATGCTGACATTATGACTCGGAATGAAT
GGGAGTAACCTCTAAAGTTGTTAACCCCTGCGTTACGACTGTAATACTGAGCCTTACTTA
A L I G D F N N W D A N A D I M T R N E 720

TTGGTGCTGGGAGATTTTCTGCCAAATAATGTGGATGGTTCCTCTGCAATTCCTCATG
AACCACAGACCCCTCTAAAAGACGGTTTATTACACCTACCAAGAGGACGTTAAGGAGTAC
F G V W E I F L P N N V D G S P A I P H 780

68/75

SnaBI

GGTCCAGAGTGAAGATACGTATGGACACTCCATCAGGTGTTAAGGATTCATTCTGCTT 840
CCAGTCTCACTTCTATGCATACCTGTGAGGTAGTCCACAATTCCTAAGGTAAGGACGAA
G S R V K I R M D T P S G V K D S I P A

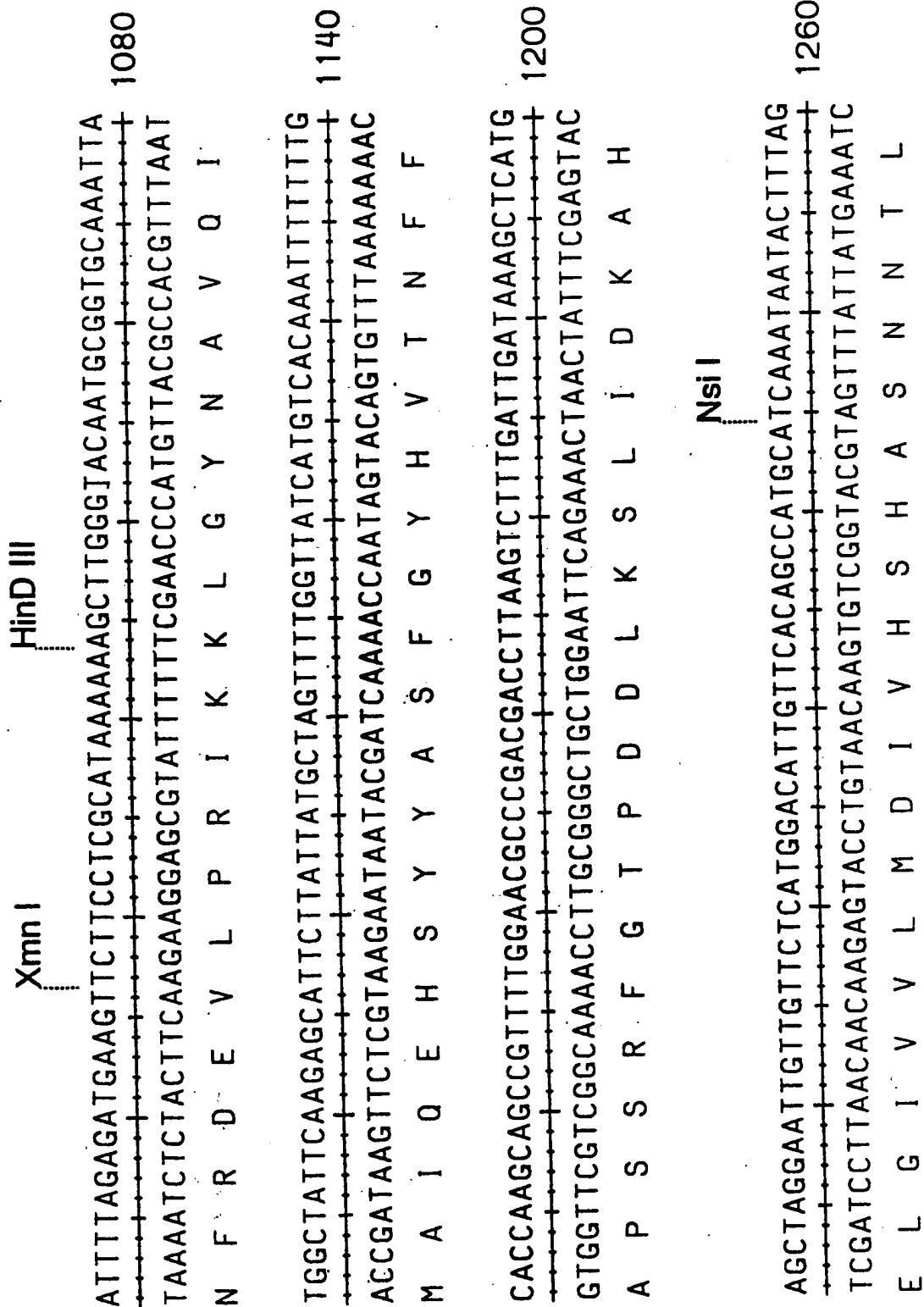
GGATCAACTACTCTTCACAGCTTCCCTGATGAAATTCATATAATGGAATATATTATGATC 900
CCTAGTTGATGAGAAGTGTCGAAGGACTACTTTAAGGTATATTACCTTATATAATACTAG
W I N Y S S Q L P D E I P Y N G I Y Y D

CACCCGAAGAGGAGGTATATCTTCCAACACCCACGGCCAAAGAAACCAAGTCGCTGA 960
GTGGGCTTCTCCTCTCCATATAGAAGGTGTGGTGCCGGTTCTTTGGTTTCAGCGACT
P P E E R Y I F Q H P R P K K S L

GAATATATGAATCTCATATTGGAATGAGTAGTCCGGAGCCCTAAATTAACATCATCGTGA 1020
CTTATATACTTAGAGTATAACCTTACTCATCAGGCCCTCGGATTTTAATTGAGTATGCACT
R I Y E S H I G M S S P E P K I N S Y V

Fig. 12
SHEET 4

69/75

Fig. 12
SHEET 5

70/75

Sac I

ATGGACTGAACATGTTTGACGGCACCAGTAGTTGTTACTTTCACTCTGGAGCTCGTGGTT
TACCTGACTTGTACAAACTGCCGTGGCTATCAACAATGAAAGTGAGACCTCGAGCACCAA
D G L N M F D G T D S C Y F H S G A R G 1320

ATCATTGGATGTGGGATTCCCGCCTTTTAACTATGGAACTGGGAGGTACTTAGGTATC
TAGTAACCTACACCCCTAAGGCGGAAAAATTGATACCTTTGACCCCTCCATGAATCCATAG
Y H W M W D S R L F N Y G N W E V L R Y 1380

TTCCTCAAATGCGAGATGGTGGTGGATGAGTTCAAATTTGATGGATTTAGATTTGATG
AAGAGAGTTTACGCTCTACCACCAACCTACTCAAGTTTAACTACCTAAATCTAAACTAG
L L S N A R W W L D E F K F D G F R F D 1440

GTGTGACATCAATGATGTATACTCACCACGGATTATCGGTGGGATTCACCTGGGAACACG
CACACTGTAGTTACTACATATGAGTGGTGCCTAATAGCCACCCCTAAGTGACCCCTTGATGC
G V T S M M Y T H H G L S V G F T G N Y 1500

Fig. 12
SHEET 6

71/75

Hinc II

AGGAATACTTTGGACTCGCAACTGATGTGGATGCTGTTGTGTATCTGATGCTGGTCAACG 1560

TCCITTATGAAACCTGAGCGTTGACTACACCTACGACACACATAGACTACGACCAGTTGC

E E Y F G L A T D V D A V V Y L M L V N

ATCTTATTCATGGGCTTTTCCAGATGCAATTACCATTTGGTGAAGATGTTAGCGGAATGC 1620

TAGAATAAGTACCCGAAAGGGTCTACGTTAATGGTAACCACTTCTACAATCGCCTTACG

D L I H G L F P D A I T I G E D V S G M

CGACATTTTGTATTCCCGTTCAAGATGGGGGTGTGGCTTTGACTATCGGCTGCATATGG 1680

GCTGTAACACATAAGGGCAAGTTCTACCCCCACACCGAAACTGATAGCCGACGTATACC

P T F C I P V Q D G G V G F D Y R L H M

CAATTGCTGATAAATGGATTGAGTTGCTCAAGAAACGGGATGAGGATTGGAGAGTGGGTG 1740

GTTAACGACTATTACCTAACTCAACGAGTTCTTTGCCCTACTCCTAACCTCTCACCCAC

A I A D K W I E L L K K R D E D W R V G

ATATTGTTTCATACACTGACAAATAGAAGATGGTCGGAAAGTGTGTTTCATACGCTGAAA 1800

TATAACAAGTATGTGACTGTTTATCTTCTACCGCCTTTTCACACAAGTATGCGACTTT

D I V H T L T N R R W S E K C V S Y A E

Fig 12
SHEET 7

72/75

Fig 12
SHEET 8

1860

GTCAATCAAGCTCTAGTCGGTGATAAAACTATAGCATCTGGCTGATGGACAAGGATA
CAGTACTAGTTCGAGATCAGCCACTATTTGATATCGTAAGACCGACTACCTGTTCCTAT
S H D Q A L V G D K T I A F W L M D K D

1920

TGATGATTTTATGGCTCTGGATAGACCGCCAACATCATTAAATAGATCGTGGGATAGCAT
ACATACTAAATACCGAGACCTATCTGGCGTTGTAGTAATTATCTAGCACCCCTATCGTA
M Y D F M A L D R P P T S L I D R G I A

Asp 718

Kpn I

1980

TGCACAAGATGATTAGGCTTGTAACCTATGGGATTAGGAGGAGAAGGTACCTAAATTTC
ACGTGTTCTACTAATCCGAACATTGATACCCCTAATCCTCCTCCTCCCATGGATTTAAAGT
L H K M I R L V T M G L G G E G Y L N F

EcoRI

2040

TGGGAAATGAATTGGCCACCCTGAGTGGATTGATTTCCCTAGGGCTGAACAACACCTCT
ACCCCTTACTTAAGCCGGTGGGACTCACCTAACTAAAGGGATCCCGACTTGTGTGGAGA
M G N E F G H P E W I D F P R A E Q H L

73/75

CTGATGACTCAGTAATTCCCGGAACCAATTCAGTTATGATAAAATGCAGACGGAGATTG 2100

GACTACTGAGTCATTAAGGGCCTTGGTTAAGTCAATACTATTACGCTGCCCTCTAAAC

S D D S V I P G N Q F S Y D K C R R R F

Ssp I

ACCTGGGAGATGCAGAATATTTAAGATACCGTGGTTGCAAGAAATTTGACCGGGCTATGC 2160

TGGACCCCTCTACGCTTATAAATTTCTATGGCACCCCAACGTTCTTAACTGGCCCGATACG

D L G D A E Y L R Y R G L Q E F D R A M

AGTATCTTGAAGATAAATATGAGTTTATGACTTCAGAACACCAGTTTCATATCACGAAAGG 2220

TCATAGAACTTCTATTATCTCAATACTGAAGTCTTGGTCAAGTATAGTGCTTTCC

Q Y L E D K Y E F M T S E H Q F I S R K

ATGAAGGAGATAGGATGTATTTGAAAAAGGAACCTAGTTTTTGTCTTTAATTTTC 2280

TACTTCCTCTATCCTACTAACATAAACTTTTTCCTTTGGATCAAAACAGAAATTAAAG

D E G D R M I V F E K G N L V F V F N F

ACTGGACAAAAGCTATTCAGACTATCGCATAGGCTGCCTGAAGCCTGGAAAATACAAGG 2340

TGACCTGTTTTTCGATAAGTCTGATAGCGTATCCGACGGACTTCGGACCTTTTATGTTCC

H W T K S Y S D Y R I G C L K P G K Y K

Fig. 12
SHEET 9

74/75

TTGCC TTGGACTCAGATCCACTTTT TGGTGGCTTCGGGAGAA TTGATCATAATG CCG 2400
AACGGAACCTGAGTCTACTAGGTGAAAAACCACCGAAGCCCTCTTA ACTAGTATTACGGC
V A L D S D D P L F G G F G R I D H N A

Ssp I

AAATTTACCTTTGAAGGATGGTATGATGATCGTCCTCGTTCAATTA TGGTGTATGCAC 2460
TTATAAAGTGGAACCTTCCTACCATACTACTAGCAGGAGCAAGTTAATACCACATACGTG
E Y F T F E G W Y D D R P R S I M V Y A

CTTGTAGAACAGCAGTGGTCTATGCACTAGTAGACAAGAAGAAGAAGAAGAAG 2520
GAACATCTTGTGTCACCAGATACGTGATCATCTGTTCTTCTTCTTCTTCTTCTTC
P C R T A V V Y A L V D K E E E E E E

AAGAAGAAGTAGCAGTAGTAGAAGAAGTAGTAGAAGAAGAAGTAGAAGCAACTTGTG 2578
TTCTTCTTCATCGTCATCATCTTCTTCTTCATCATCATCTTCTTCTTCTTCTTGAACAC
E E E V A V V E E V V V E E E

Fig 12
SHEET 10

75/75

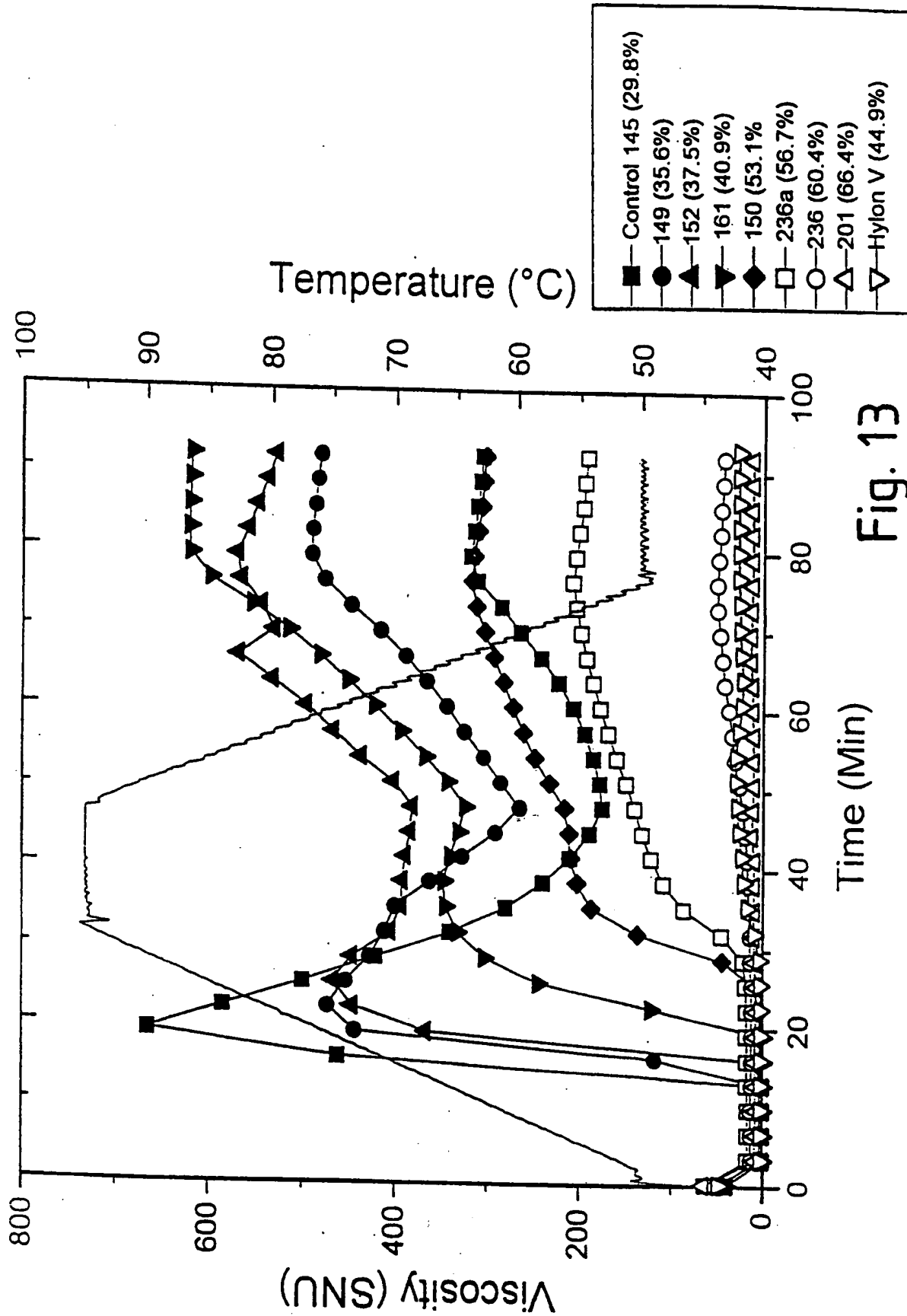


Fig. 13